1994

Large deviation results for branching processes in fixed and random environments

Anand N. Vidyashankar
Iowa State University

Order Number 9503602

Large deviation results for branching processes in fixed and random environments

Vidyashankar, Anand N., Ph.D.

Iowa State University, 1994

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

Large deviation results for branching processes in fixed and random environments

by

Anand N. Vidyashankar

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Departments: Statistics
Mathematics
Majors: Statistics
Mathematics

Approved:

Signature was redacted for privacy.

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa
1994

TABLE OF CONTENTS

GENERAL INTRODUCTION

Branching processes form one of the classical fields of applied probability and remain an active area of research and application. In recent years, there has been renewed interest in the old topic of large deviations, namely, the asymptotic computation of small probabilities on an exponential scale. The reasons are two-fold. On the one hand, starting with Donsker and Varadhan, a general foundation was laid that allowed one to use "general" tricks that seemed to work in several diverse situations. On the other hand, large deviation estimates have proved to be a crucial tool for studying many questions in statistics, engineering, statistical mechanics, and partial differential equations.

The large deviation questions in branching processes have a different flavor. The "general" trick from large deviation theory does not seem to work here. Also, the large deviation questions (for the mean) tie up with some of the classical open questions in the theory of branching processes. The large deviation questions have non-trivial applications to the study of algorithms in theoretical computer science (see [4] and [5]). For these reasons we study the large deviation questions that arise in branching models.

In the first paper, we study the rate of decay of the probabilities of deviation of an estimator of mean from the true mean. We show that this rate is geometric. These results depend on the known results about the rate of decay of iterates of one-dimensional generating functions.

A more detailed study of single type processes was carried out by Athreya (see [1]). He showed that the geometric rate of decay of large deviation probabilities is intimately connected to the sample paths that grow slowly. He further established that if the branching process grows "fast enough" then the large deviation probabilities decay exponentially fast. Finally, he considered the problem of decay rates for martingales and showed that this rate is super-geometric.

When one considers multitype branching processes, new questions about the rate of convergence of proportions arise. As in the single type case, we reduce the problem to the study of decay rates of iterates of two-dimensional generating functions. This problem has not been settled in the literature and we develop techniques for studying this problem. It turns out that even in this case, the rate of decay of the probabilities of large deviations and the rate of decay of the iterates of analytic functions are the same and are geometric (under further conditions). Further, we provide partial solutions to the large deviation questions for multitype processes that grow exponentially fast. We also consider rate of convergence of the martingales and show that this rate is super geometric. All of these results are contained in the second paper.

In the fourth paper of the thesis we move on to consider continuous time branching processes. In this case the situation completely changes depending on the tails of the life-time distributions $G(\cdot)$. We show that in the Malthusian Case the rate of decay of generating functionals is exponential, in the sub-exponential case the rate is $(1 - G(t))$ and in the super-exponential case the rate is super-exponential. These decay rates transform to decay rates for large deviation probabilities under appropriate moment and regularity conditions. Finally, we consider the decay rates of martingales and show that this rate is superexponential.

Having studied the rates of decay of the sample quantities from the true values, we consider large deviations that arise by considering the tails of the martingale limit W in a multitype branching process. More precisely, we consider rates of decay of $\log P(W \leq x)$ as $x \to 0$ and $\log P(W > x)$ as $x \to \infty$. The study of $\log P(W \leq x)$ is carried out by studying the density w of W near the origin. For this reason, we obtain bounds on the density near the origin and use that to extract the exponential decrease of $P(W \leq x)$. Next we consider the decay rates for $P(W > x)$ as $x \to \infty$ for a finite multitype branching process. For this reason, we develop multitype versions of the Harris function, study its properties, and use that to extract the exponential rate. The results for single type processes are available in [2] and [3]. These results are contained in the third paper.

Finally, we move on to study branching processes in random environments with an embedded random walk. In this situation, by "large deviations" we mean "by how much the large deviations of the mean measure carry over to the sequence $Z^{(n)}$." This study leads to questions about the non-degeneracy of the limit of a certain martingale sequence. This is carried out in the fifth paper while in the sixth paper the large deviations of $Z^{(n)}$ are investigated.

Each of the papers contains a list of unsolved problems. There is also a collection of statistical questions that are as yet unresolved and require further work.

Explanation of the Dissertation Format.

This thesis contains six papers that have been submitted to a scholarly journal for publication. Following the last paper is a general summary followed by a list of references cited in the general introduction.

LARGE DEVIATION RESULTS FOR BRANCHING PROCESSES

A PAPER INCLUDED IN A FESTSCHRIFT IN HONOUR OF GOPINATH KALLIANPUR,
SPRINGER-VERLAG, NEW YORK.
KRISHNA B. ATHREYA AND ANAND N. VIDYASHANKAR
IOWA STATE UNIVERSITY

1. Introduction.

Let $\{Z_n\}_0^\infty$ be a Galton-Watson branching process with offspring probability distribution $\{p_j : j = 0_1 1, 2 \ldots\}$. Assume $p_0 = 0, 1 < m \equiv \sum j p_j < \infty$ and $\sigma^2 \equiv \sum j^2 p_j - m^2 > 0$. It is known that (see Athreya & Ney [2] pp. 24) if $P(Z_0 > 0) = 1$ then $P(Z_n \to \infty) = 1$ and $Z_{n+1} Z_n^{-1} \to m$ w.p.1. Also by the central limit theorem for random sums $(Z_{n+1} Z_n^{-1} - m)\sqrt{m^n}$ converges in distribution to a mixture of normal distributions provided $0 < \sigma^2 < \infty$. There is a law of the iterated logarithm as well. The problems of large deviations of $Z_{n+1} Z_n^{-1} - m$ have not been treated in the literature and this note fills this gap. The next section treats the one dimensional case. Extension to the multitype case is outlined (without proof) in section 3.

2. The Main Result.

In this section we prove the following:

Theorem 1. *Let the offspring distribution $\{p_j\}$ satisfy $p_0 = 0$ and $p_1 \neq 0$. Let A be a Borel set in R such that for random variables $\{X_i\}_1^\infty$ that are independent with distribution $\{p_j\}$ there exists a λ in (0,1) such that*

$$G(n, A) \equiv P(\overline{X}_n - m \in A) = O(\lambda^n) \text{ as } n \to \infty \tag{1}$$

where $\overline{X}_n = n^{-1} \sum_1^n X_i$ *and* $m = \sum j p_j$. *Then, if* $P(Z_0 = 1) = 1$

(2) $$p_1^{-n} P((Z_{n+1} Z_n^{-1} - m) \in A) \to \sum_{j=1}^{\infty} G((j, A) q_j < \infty$$

where $\{q_j\}$ *is defined by its generating function* $Q(s) \equiv \sum_1^{\infty} q_j s^j$, *as the unique solution of the functional equation*

(3) $$\begin{cases} Q(f(s)) & = p_1 Q(s) \text{ for } 0 \leq s < 1 \\ Q(0) & = 0, \end{cases}$$

where $f(s) = \sum s^j p_j$

Proof. Let $f_n(s)$ be the probability generating function of Z_n with $P(Z_0 = 1) = 1$. Then it is known that (see Athreya & Ney [2] pp. 38)

(4) $$\lim_n p_1^{-n} f_n(s) \equiv Q(s) < \infty \text{ exists for } 0 \leq s < 1$$

and $Q(\cdot)$ is the unique solution of the functional equation (3).

Now, by conditioning on $Z_0, Z_1 \dots Z_n$ and using the branching property, we see that

(5) $$P\left((Z_{n+1} Z_n^{-1} - m) \in A\right) = \sum_j G(j, A) P(Z_n = j)$$

Let $h_n(j) = G(j, A) P(Z_n = j) p_1^{-n}$. By hypothesis there exists constants $C \in (0, \infty)$ and $\lambda \in (0, 1)$ such that $G(j, A) \leq C \lambda^j$ for all $j \geq 1$.

Let $r_n(j) = C \lambda^j P(Z_n = j) p_1^{-n}$

Then by (4)

$$\sum_j r_n(j) = C p_1^{-n} f_n(\lambda) \to C Q(\lambda) < \infty$$

Also we have, for $j = 1, 2, \dots$.

$$0 \leq h_n(j) \leq r_n(j)$$

and

$$h_n(j) \to G(j, A)q_j.$$

(This last assertion follows from (4)).

So by a slight generalization of the dominated convergence theorem (see Royden [5] pp. 92, pp. 270) applied to the counting measure space on the positive integers we get from (5),

$$p_1^{-n} P((Z_{n+1} Z_n^{-1} - m) \in A) \to \sum_j G(j, A) q_j < \infty. \quad \square$$

Corollary 1. *Let $\{p_j\}$ satisfy $f(\theta_0) < \infty$ for some $1 < \theta_0 < \infty$ and $p_0 = 0, 0 < p_1 < 1$. Let A be any Borel set $\subset R - (-\epsilon, \epsilon)$ for some $\epsilon > 0$. Then (2) holds.*

Proof.

$$\begin{aligned} G(n, A) &\le P(|\overline{X}_n - m| > \epsilon) \\ &\le P(\overline{X}_n > m + \epsilon) + P(\overline{X}_n < m - \epsilon) \\ &\le P(\alpha^{S_n} > \alpha^{n(m+\epsilon)}) + P(\beta^{S_n} > \beta^{n(m-\epsilon)}) \end{aligned}$$

where $S_n = \sum_1^n X_i$, α and β are arbitrary constants in $(1, \theta_0)$ and (0,1) respectively. Thus, by Markov's inequality

$$G(n, A) \le (f(\alpha)\alpha^{-(m+\epsilon)})^n + (f(\beta)\beta^{-(m-\epsilon)})^n.$$

It can be verified that for every $0 < \epsilon < 1$ there exists α_0 in $(1, \theta_0)$ and β_0 in (0,1) such that

$$0 < f(\alpha_0)\alpha_0^{-(m+\epsilon)} < 1 \text{ and } 0 < f(\beta_0)\beta_0^{-(m-\epsilon)} < 1.$$

This yields (1) and so (2) follows.

Corollary 2. *Let $\{p_j\}$ satisfy $f(\theta_0) < \infty$ for some $1 < \theta_0 < \infty$. Let $p_0 > 0, p_1 > 0$ and $1 < m = \sum jp_j < \infty$. Let $P(Z_0 = 1) = 1$. Then for every Borel set $A \subset R - (-\epsilon, \epsilon)$ for some $\epsilon > 0$*

$$\gamma^{-n} P((Z_{n+1}Z_n^{-1} - m) \in A | Z_n > 0) \to \sum_j G(j, A) \left(\frac{q_j}{1-q} \right) \tag{6}$$

where $\gamma = f'(q), q$ the smallest root of $s = f(s)$ in [0,1], $Q(s) \equiv \sum_0^\infty q_j s^j$, the unique solution of the equation

$$\begin{cases} Q(f(s)) & = \gamma Q(s) \qquad 0 \le s < 1 \\ Q(0) & = 0 \end{cases}$$

Proof. As in the proof of Theorem 1

$$\begin{aligned} &P((Z_{n+1}Z_n^{-1} - m) \in A | Z_n > 0) \\ &\qquad = \sum_{j=1}^\infty G(j, A) P(Z_n = j | Z_n > 0). \end{aligned}$$

It is known that (see Athreya and Ney [2] pp. 40) if $P(Z_0 = 1) = 1, p_0 > 0, 1 < m < \infty$ then $\gamma^{-n} P(Z_n = j | Z_n > 0) \to \frac{q_j}{(1-q)}$. The rest of the argument is the same as in Theorem 1 and Corollary 1.

Remark 1. The conclusion (2) is a stronger form of the usual large deviation type result involving rate functions for the exponential decay. Indeed if (1) holds for a Borel set $A \subset R$ then

$$\frac{1}{n} \log P((Z_{n+1}Z_n^{-1} - m) \in A) \to \log p_1. \tag{7}$$

In particular, if $f(\theta_0) < \infty$ for some $1 < \theta_0 < \infty$ then by Corollary 1, (7) above holds for all Borel sets $A \subset R - (-\epsilon, \epsilon)$ for some $\epsilon > 0$. Thus if a rate function $I(\cdot)$ were to exist that satisfies the usual conditions for being called a *good rate function*, (see Deuschel & Stroock [3]) then by considering an open interval $(x_0 - h, x_0 + h)$ and a closed interval $[x_0 - h, x_0 + h]$ in $(0,\infty)\cup(-\infty,0)$ we see that

$$-\liminf_{x_0-h<x<x_0+h} I(x) \le \log p_1$$

and

$$-\liminf_{x_0-h\le x\le x_0+h} \ge \log p_1.$$

Since $I(\cdot)$ is lower semicontinuous on R it must follow that $I(x) = -\log p_1$ for all $x \neq 0$. On the other hand if $A = (0,\infty)$ and if $\sum j^2 p_j < \infty$ (and in particular if $f(\theta_0) < \infty$ for some $1 < \theta_0 < \infty$) then by the central limit theorem (see Athreya and Ney [2] pp. 55)

$$P((Z_{n+1}Z_n^{-1} - m) \in A) \to \frac{1}{2}$$

and hence

$$\frac{1}{n}\log P((Z_{n+1}Z_n^{-1} - m) \in A) \to 0.$$

This is also true for $A = [0,\infty)$. Thus, the rate function must satisfy

$$0 \le \inf_{x\ge 0} I(x) \le \inf_{x>0} I(x) \le 0.$$

But $\inf_{x>0} I(x) = -\log p_1$ and so we have a contradiction. This shows that there is no good rate function for the large deviation problem on hand. Nevertheless (7) does hold for all A satisfying (1) and in particular, for all Borel sets $A \subset R - (-\epsilon,\epsilon)$ for some $\epsilon > 0$.

Remark 2. Regarding complete convergence, that is, convergence of the series $\sum_n P(Z_{n+1}Z_n^{-1} - m \in A)$ we note that if $\sum j^2 p_j < \infty$ and $A \subset R - (-\epsilon,\epsilon)$ for some $\epsilon > 0$ then by the Erdos-Hsu Robbins theorem (see [4])

$$\sum G(j, A) < \infty$$

where G is as in (1).

Thus

$$\sum_n P(Z_{n+1}Z_n^{-1} - m \in A) = \sum_n EG(Z_n, A) = E\left(\sum_n G(Z_n, A)\right)$$
$$\leq \sum_j G(j, A) < \infty$$

We conjecture that $EG(Z_n, A) < \infty$ just with $1 < m < \infty$. (See also Asmussen and Kurtz [1]).

Remark 3. (Open problem) The assumption that (1) holds could be too strong for (2). From (5) we see that

$$\underline{\lim} p_1^{-n} P(Z_{n+1}Z_n^{-1} - m \in A) \geq \sum_j G(j, A) q_j. \tag{8}$$

At the moment not much is known about $\{q_j\}$. An interesting open problem is to investigate the growth rate of $\{q_j\}$ and relate that to the convergence of $\sum_j G(j, A) q_j$ as well as improving (8) to a full convergence result.

3. Extensions to multitype case.

Let $\{Z_n\}_0^\infty$ be a p-type ($p > 1$), positively regular, supercritical branching process with offspring generating functions $f^{(i)}(s)$ for $i = 1, 2 \ldots p$. Assume $f^{(i)}(0) = 0$ for all i. Let $m_{ij} = \frac{\partial f^{(i)}}{\partial s_j}(1)$ where $\underline{1} = (1, 1, \ldots, 1)$ and $a_{ij} = \frac{\partial f^{(i)}}{\partial s_j}(0)$. Let γ be the Perron-Frobenius root of the matrix $A = ((a_{ij}))$ which is assumed to be positively regular. Let ρ be the maximal eigenvalue of M with normalized eigenvectors u and v such that $u'M = \rho u', Mu = \rho u, u \cdot \underline{1} = \underline{1}, u \cdot v = \underline{1}$. Let $f_1(s) = (f^{(1)}(s), f^{(2)}(s), \ldots, f^{(p)}(s))$ be a map of the unit cube $C = \{s : s = (s_1, s_2, \ldots s_p), 0 \leq s_i \leq 1\}$ onto itself and $f_n(\cdot)$ be its nth iterate. Then the following results hold.

Theorem 2. *There exists a map Q of the open unit cube C to R_+^p such that*

$$\frac{f_n(s)}{\gamma^n} \to Q(s)$$

and Q is the unique solution of

$$Q(f(s)) = \gamma Q(s)$$

$$Q(0) = 0$$

Theorem 3. *Assume $f^{(i)}(\theta_0 1) < \infty$ for some $1 < \theta_0 < \infty$ and for all $1 \leq i \leq p$. Let $P(1 \cdot Z_0 = 1) = \underline{1}$. Then for any vector ℓ and $\epsilon > 0$, $\lim_n \frac{1}{\gamma^n} P(\frac{\ell \cdot Z_{n+1}}{\underline{1} \cdot Z_n} - \frac{\ell \cdot Z_n M}{\underline{1} \cdot Z_n} > \epsilon)$ and $\lim_n \frac{1}{\gamma^n} P(\frac{\ell \cdot Z_n}{\underline{1} \cdot Z_n} > \ell \cdot v + \epsilon)$ exist and are finite and positive.*

Theorem 2 which is a p-type extension of Theorem 1 (p. 40 of Athreya and Ney [2]) is new. The proof of Theorem 3 uses Theorem 2 as in the proof of Theorem 1 although the arguments are more involved. The proofs of both these results will be given elsewhere. The problem of obtaining the conclusion of Theorem 3 without the exponential moment hypothesis is an interesting open problem.

References.

1. Asmussen, and Kurtz, G., *Necessary and Sufficient Conditions for Compute Convergence in the Law of Large Numbers*, Annals of Probability (1980), 176-182.
2. Athreya, K.B. and Ney, P.E., *Branching Processes*, Springer-Verlag, Berlin, 1972.
3. Deuschel, J.-D. and Stroock, D.W., *Large Deviations*, Academic Press, N.Y., 1989.
4. Hsu, P.L. and Robbins, H., *Complete Convergence and the Law of Large Numbers*, Proc. Natl. Acad. Sci. U.S.A. **33** (1947), 25-31.
5. Royden, H.L., *Real Analysis*, Macmillan Publishing Company, N.Y., 1987, Third Edition.

LARGE DEVIATION RATES FOR BRANCHING PROCESSES - II, THE MULTITYPE CASE

A PAPER SUBMITTED TO ANNALS OF APPLIED PROBABILITY
KRISHNA B. ATHREYA AND ANAND N. VIDYASHANKAR

Departments of Mathematics and Statistics
Iowa State University
Ames, Iowa 50011

Abstract. Let $\{Z_n: n \geq 0\}$ be a p-type ($p \geq 2$) supercritical branching process with mean matrix M. It is known that for any ℓ in R^p, $\left(\frac{\ell \cdot Z_n}{1 \cdot Z_n} - \frac{\ell \cdot (Z_n M)}{1 \cdot Z_n}\right)$ and $\left(\frac{\ell \cdot Z_n}{1 \cdot Z_n} - \frac{\ell \cdot v^{(1)}}{1 \cdot v^{(1)}}\right)$ converge to 0 with probability one on the set of non-extinction where $v^{(1)}$ is the left eigenvector of M corresponding to its maximal eigenvalue ρ. In this paper we study the large deviation aspects of this convergence. It is shown that the large deviation probabilities for these two sequences decay geometrically and under appropriate conditioning supergeometrically.

1. Introduction. Let $\{Z_n: n \geq 0\}$ be a supercritical p-type Galton-Watson branching process (see [2] for definition) with offspring generating functions $f^{(i)}(s)\, i = 1, 2, \ldots, p$ and mean matrix M. Let ρ be the maximal eigenvalue of M (necessarily greater than 1) with the corresponding left and right eigenvectors $v^{(1)}$ and $u^{(1)}$ respectively. It is known that for any vector ℓ (see [2])

$$\left(\frac{\ell \cdot Z_{n+1}}{1 \cdot Z_n} - \frac{\ell \cdot (Z_n M)}{1 \cdot Z_n}\right) \text{ and } \left(\frac{\ell \cdot Z_n}{1 \cdot Z_n} - \frac{\ell \cdot v^{(1)}}{1 \cdot v^{(1)}}\right)$$

converge to 0 with probability 1 (wp1) on the set of non-extinction and that $\{W_n \equiv \frac{u^{(1)} \cdot Z_n}{\rho^n} : n \geq 0\}$ is a nonnegative martingale sequence and hence converges to a non-negative random variable W wp1.

The questions addressed in this paper concern the large deviation aspects of the above convergence. It is shown that, under certain moment conditions, the rate of decay of the probabilities of large deviations is geometric, while conditionally on $W \geq a$ $(a > 0)$, the rate is supergeometric. The corresponding results for single type branching process are available in [1] and [3]. In [8] large deviation aspects of $P(W \leq x)$ and $P(W \geq x)$ as $x \to 0$ and $x \to \infty$ respectively are studied.

As in those papers, we reduce the problem (using the moment conditions on the offspring distributions) to a study of decay rates of iterates of generating functions $f^{(i)}$.

The paper is organized as follows: Section 2 contains notations, definitions, and assumptions; Section 3 gives statements of the results; and Section 4 contains the proofs. Section 5 is devoted to some open problems. For ease of exposition we assume $p = 2$ throughout the rest of the paper.

2. Notations, Definitions, and Assumptions.

(1) $\mathcal{C}_2 = [0,1] \times [0,1]$ is the unit square in R^2, the two dimensional Euclidean space.

(2) $\mathcal{A}_2 = \{(i_1, i_2)\colon\ i, \in \mathbb{Z}_+, i_2 \in \mathbb{Z}_+\}$ where $\mathbb{Z}_+$ is the set of all non-negative integers.

(3) For $s \in \mathcal{C}_2$ and $j \in \mathcal{A}_2$, $\quad s^j = s_1^{j_1} s_2^{j_2}$

(4) $1 = (1,1) \quad e_1 = (1,0), \quad e_2 = (0,1), \quad 0 = (0,0)$

(5) $Z_n = (Z_n^{(1)}, Z_n^{(2)})$ is the population vector of the nth generation.

(6) $P_i(\cdot) = P(\cdot|Z_0 = e_i)$ and $E_i(\cdot) = E(\cdot|Z_0 = e_i)$ for $i = 1, 2$.

(7) $P_i(j_1, j_2) = P(Z_1 = (j_1, j_2)|Z_0 = e_i)$

(8) For $s \in \mathcal{C}_2, f_i^{(n)}(s) = E(s^{Z_n}|Z_0 = e_i), i = 1, 2$. If $n = 1$, we shall write

$f_1(s)$ and $f_2(s)$ for $f_1^{(1)}(s)$ and $f_2^{(1)}(s)$.

(9) For $n \geq 0$ and $s \in \mathcal{C}_2, f^{(n)}(s) = (f_1^{(n)}(s), f_2^{(n)}(s))$ where for $n = 0, f_1^{(0)}(s) \triangleq s$, and $f(s) \equiv f^{(1)}(s)$. It is known that (see [2]) for all $n \geq 1$,

$$f^{(n)}(s) = f(f^{(n-1)}(s)).$$

(10) For $s \in \mathcal{C}_2, ||s|| = \max(s_1, s_2)$ and $||E(\cdot)|| = \max(|E_1(\cdot)|, |E_2(\cdot)|)$.

(11) $R_+^2 = \{(x_1, x_2) : x_1 \geq 0, \quad x_2 \geq 0\}$.

(12) For $i, j = 1, 2, D_{ij}(s) = \dfrac{\partial f_i(s)}{\partial sj}$ and $a_{ij} = D_{ij}(0)$.

(13) $M = (m_{ij})$ where $m_{ij} = \dfrac{\partial f_i}{\partial s_j}(1)$ for $i, j = 1, 2$.

(14) For any matrix E, its transpose will be denoted by E^t.

Assumptions

A1 $f(0,0) = 0$.

A2 M is positively regular with maximum eigenvalue ρ and the associated right and left eigenvectors $u^{(1)}$ and $v^{(1)}$ respectively.

A3 $\rho > 1$.

A4 A is positively regular with maximum eigenvalue γ and the associated right and left eigenvectors $u^{(0)}$ and $v^{(0)}$ respectively. The row sums of A are strictly less than 1. (It follows that $0 < \gamma < 1$.)

A5 $||E(e^{\theta_0(1 \cdot Z_1)})|| < \infty$ for some $\theta_0 > 0$.

A6 $||E(1 \cdot Z_1)^{2r_0+\delta}|| < \infty$ where r_0 is such that $(m_{21} + m_{22})^{r_0}\gamma > 1$ and $\delta > 0$.

A7 $P_i(Z_i^{(1)} \leq 1) = 0$ and $P_i(Z_i^{(1)} = 2) > 0$ for $i = 1, 2$.

3. Statements of Results.

It is known (see [2]) that $f^{(n)}(s) \to 0$ (for $s \neq 1$) as $n \to \infty$. Our first theorem gives the rate of convergence under A4.

Theorem 1. *Under A1 and A4, there exists a map $Q: \mathcal{C}_2 \to R_+^2$ such that*

$$\frac{f_n(s)}{\gamma^n} \to Q(s) \text{ as } n \to \infty$$

and $Q(\cdot)$ satisfies the vector functional equation

$$Q(f(s)) = \gamma Q(s)$$

and the constraints

$$Q(0) = 0$$

$$\lim_{s \uparrow 1} Q(s) = \infty$$

$$\text{and } 0 < Q(s) < \infty \qquad \text{for } 0 < s < 1$$

Remark 1. The assumption A1 can be removed by considering $f_n(s) - q$ where q is the extinction probability vector, i.e. $q_i = P_i(Z_n = 0 \text{ for some } n \geq 1)$ and $A = (a_{ij})$ where $a_{ij} = D_{ij}(q)$.

Theorem 2. *Under A1 and A4 the function $Q(\cdot)$ defined in Theorem 1 is the unique solution of the functional equation*

$$Q(f(s)) = \gamma Q(s) \qquad \text{for } s \neq 1$$

satisfying

$$Q(0) = 0 \text{ and } \lim_{s \to 0} Q'(s) = P_0, \text{ where } P_0 = \lim_{n \to \infty} \gamma^{-n} A^n$$

(which exists by the Perron-Frobenius theorem).

Remark 2. The next theorem gives a rate of decay for the generating functions when $A = 0$ and every particle produces at least two particles of its kind.

Theorem 3. *Under A7*

$$\lim_{n \to \infty} \frac{\log f^{(n)}(s)}{2^n} \equiv R_i(s) \qquad \text{for } i = 1, 2$$

exists and satisfies the vector functional equation

$$R_i(f(s)) = 2R_i(s)$$

$$\text{and } \lim_{s \downarrow 0} R_i(s) = -\infty$$

Theorem 4 is the large deviation theorem for functionals of the process under exponential moment hypothesis on the offspring distribution function.

Theorem 4. *Assume that A1– A5 holds. Then for $\ell \neq 0$, and $\varepsilon > 0$*

$$\lim_{n\to\infty} \gamma^{-n} P_i \left(\left| \frac{\ell \cdot Z_{n+1}}{1 \cdot Z_n} - \frac{\ell \cdot (Z_n M)}{1 \cdot Z_n} \right| > \varepsilon \right)$$
$$\lim_{n\to\infty} \gamma^{-n} P_i \left(\left| \frac{\ell \cdot Z_n}{1 \cdot Z_n} - \frac{\ell \cdot v^{(1)}}{1 \cdot v^{(1)}} \right| > \varepsilon \right)$$

exist and are positive and finite for $i = 1, 2$.

The next theorem weakens the exponential moment hypothesis to polynomial moment hypothesis.

Theorem 5. *Assume A1 – A4 and A6 hold. Then the conclusions of Theorem 4 are valid.*

Our next theorem considers the case when A4 does not hold. In this case the rate of decay of probabilities of large deviations is super-geometric.

Theorem 6. *Assume A1 – A3, A5 and A7 hold. Then for $\ell \neq 0$, and $\varepsilon > 0$, there exists constants $0 < C_1(\varepsilon), C_2(\varepsilon) < \infty$ and $0 < \lambda_1(\varepsilon), \lambda_2(\varepsilon) < 1$ such that*

$$P_i \left(\left| \frac{\ell \cdot Z_{n+1}}{1 \cdot Z_n} - \frac{\ell \cdot (Z_n M)}{1 \cdot Z_n} \right| \geq \varepsilon \right) \leq C_1 \cdot \lambda_1^{(2^n)}$$
$$\text{and} \quad P_i \left(\left| \frac{\ell \cdot Z_n}{1 \cdot Z_n} - \frac{\ell \cdot v^{(1)}}{1 \cdot v^{(1)}} \right| \geq \varepsilon \right) \leq C_2 \lambda_2^{(2^n)}$$

for $i = 1, 2$.

Theorem 7 is needed in the proof of Theorem 8.

Theorem 7. *Under A5 there exists $\theta_0 > 0$ such that*

$$\sup_{n \geq 1} \| E(e^{\theta_0 W_n}) \| < \infty.$$

Theorem 8 asserts that the decay rate of $P(|W - W_n| \geq \varepsilon)$ is always supergeometric.

Theorem 8. *Assume A1 – A3 and A5 hold. Then, there exist constants* $0 < C_3 < \infty$ *and* $0 < \lambda_3 < \infty$ *such that for* $\varepsilon > 0$

$$P(|W - W_n| \geq \varepsilon) \leq C_3 e^{-\lambda_3 \varepsilon^{2/3}(\rho^{1/3})^n}$$

The next theorem shows that conditioned on $W \geq a, a > 0$, *large deviation probabilities in Theorem 6 decay supergeometrically.*

Theorem 9. *Assume A1 - A3 and A5 hold. Then there exist constants* $0 < C_4, C_5, C_6, C_7 < \infty$ *and* $\lambda_4, \lambda_5 > 0$ *such that for every* $\varepsilon > 0$ *and* $a > 0$ *there exists* $0 < I(\varepsilon) < \infty$ *such that*

$$P_i\left(|\frac{\ell \cdot Z_{n+1}}{1 \cdot Z_n} - \frac{\ell \cdot (Z_n M)}{1 \cdot Z_n}| > \varepsilon | W \geq a\right) \leq C_4 e^{-aI(\varepsilon)\cdot\xi\rho^n} + C_5 e^{-\lambda_4 (a(1-\xi))^{2/3}(\rho^{2/3})^n}$$

and

$$P_i\left(|\frac{\ell \cdot Z_n}{1 \cdot Z_n} - \frac{\ell \cdot v^{(1)}}{1 \cdot v^{(1)}}| > \varepsilon | W \geq a\right) \leq C_6 e^{-aI(\varepsilon)\xi\rho^n} + C_7 e^{-\lambda_5 (a(1-\xi))^{2/3}(\rho^{2/3})^n}$$

for every $0 < \xi < 1$.

4. Proofs.

Proof of Theorem 1. Observe that for $s \in C_2$

$$f(s) = sB + g(s)$$

$$\text{where } B = A^t \text{ and } g(s) = (g_1(s), g_2(s))$$

$$\text{where } g_i(s) = \sum_{j_1 \geq 1} \sum_{j_2 \geq 1} P_i(j_1, j_2) s_1^{j_1} s_2^{j_2} + \sum_{j_1 \geq 2} P_i(j_1, 0) s_1^{j_1} + \sum_{j_2 \geq 2} P_i(0, j_2) s_2^{j_2}$$

Iterating the above equation we have

$$f^{(n)}(s) = sB^n + \sum_{k=0}^{(n-1)} g(f^{(n-1-k)}(s))B^k$$

Hence

$$\gamma^{-n} f^{(n)}(s) = \gamma^{-n} s B^n + \sum_{k=0}^{n-1} \gamma^{-n} g(f^{(n-1-k)}(s)) B^k \tag{1}$$

the first term converges to sP_0^t (by the Perron-Frobenius theorem) as $n \to \infty$.

The second term is the same as

$$\gamma^{-1} \left[\sum_{k=0}^{n-1} \gamma^{-k} g(f^{(k)}(s))(\gamma^{-1}B)^{n-1-k} \right].$$

We shall show that

$$\sum_{k \geq 0} \frac{g(f^{(k)}(s))}{\gamma^k} < \infty. \tag{2}$$

From this it will follow (by the dominated convergence theorem applied to counting measure space $\mathcal{A}_2$) (see [1] and [7]) that

$$\lim_{n \to \infty} \frac{f^{(n)}(s)}{\gamma^n} = sP_0^t + \left(\sum_{n \geq 0} \frac{g(f^{(n)}(s))}{\gamma^n} \right) P_0^t \equiv Q(s), \tag{3}$$

Note that

$$||g(s)|| \leq C||s||^2 \text{ for } s \in C_2. \tag{4}$$

Using (3) it is easy to see that, to establish (2), it is enough to establish the following:

(a) $\{\frac{f_n(s)}{\gamma^n} : n \geq 0\}$ is a bounded sequence for $s \in \mathcal{C}_2 - 1$.

(b) $||f_n(s)|| \leq C\delta^n$ for $s \in \mathcal{C}_2 - 1$ where C is a finite positive constant and $0 < \delta < 1$, (both C & δ depend on s).

Lemma 1 below is a real variable lemma needed for establishing (a) and may also be of independent interest. Lemma 2 contains the proof of (b).

□

Lemma 1. *Let $\{a_n : n \geq 0\}$ be a sequence of positive numbers satisfying*

$$a_n \leq C_1 + C_2 \sum_{k=0}^{n-1} \eta^k a_k \quad n \geq 1 \tag{5}$$

where $0 < \eta < 1$ and C_1 and C_2 are finite positive constants. Then

$$\sup_{n \geq 0} a_n < \infty.$$

Proof. Note that

$$\begin{aligned} & a_1 \leq C_1 + C_2 a_0 \\ \text{and} \quad & a_2 \leq (C_1 + C_2 a_0)(1 + C_2 \eta). \\ \text{Iteration yields} \quad & a_n \leq (C_1 + C_2 a_0) \prod_{j=1}^{n-1} (1 + C_2 \eta^j) \\ & \leq (C_1 + C_2 a_0) e^{C_2 \eta/(1-\eta)} < \infty \end{aligned}$$

using the simple bound $1 + x \leq e^x$ for $x \geq 0$.

Lemma 2. *For each $||s|| \leq 1, s \neq 1$, there exist positive constants N, C, and δ (depending on s) with $N \geq 1, 0 < \delta < 1$, such that $\forall n \geq 1$*

$$||f^{n+N}(s)|| \leq C\delta^n \quad \text{for } s \neq 1. \tag{6}$$

Proof. Since $f_n(s) \to 0$, given $\varepsilon > 0, \exists\ N_0(\varepsilon, s)$ such that

$$||f_n(s)|| < \varepsilon \quad \forall n \geq N_0, \quad \text{for } s \neq 1.$$

Also recalling that

$$0 < \sum_{j=1}^{2} a_{ij} < 1 \quad \text{for all } 1 \leq i \leq 2,$$

we see by the continuity of $\sum_{j=1}^{2} D_{ij}(s), \exists\ \eta > 0$ and $0 < \delta < 1$ such that for $||s|| < \eta$

$$\sum_{j=1}^{2} D_{ij}(s) \leq \delta \quad \text{for all } i = 1, 2.$$

For $n \geq 1$

$$\begin{aligned} ||f^{(n+N_0)}(s)|| &= ||f^{(n+N_0)}(s) - f^{(n+N_0)}(0)|| \\ &= \max_{1 \leq i \leq 2} (f_i(f^{(n+N_0-1)}(s)) - f_i(f^{(n+N_0-1)}(s))). \end{aligned}$$

But by the mean value theorem,

$$\begin{aligned} f_i(f^{(n+N_0-1)}(s)) - f_i(f^{(n+N_0-1)}(0)) &= \sum_{j=1}^{2} f_j^{(n+N_0-1)}(s) \frac{\partial f_i}{\partial s_j}(s^*) \text{where } s^* \in (0, f^{(n+N_0-1)}(s)) \\ &\leq ||f^{(n+N_0-1)}(s)|| \sum_{j=1}^{2} D_{ij}(s^*) \quad (\text{ with } \varepsilon = \eta) \\ &\leq ||f^{(n+N_0-1)}(s)|| \delta. \end{aligned}$$

Iterating we have

$$\begin{aligned} ||f^{(n+N_0)}(s)|| &\leq ||f^{(N_0)}(s)|| \delta^n \\ &\leq C \delta^n \end{aligned}$$

completing the proof of the lemma.

Now to complete the proof of (a) (and hence Theorem 1) observe that by (1), (4), lemma 2 and Perron-Frobenius Theorem, $a_n \equiv ||f^{(n)}(s)/\gamma^n||$ satisfies the hypothesis of lemma 1 for an appropriate choice of C_1, C_2 and η. Finally, the facts that Q is non-trivial and satisfies the functional equation follow from (3).

□

Remark 3. Even for the single type case one can construct a proof based on the above method. However in this case, finiteness of $\sum_{n \geq 1} \frac{g(f^{(n)}(s))}{\gamma^n}$ (see (2)) where

$\gamma = f'(0)$ can be seen by an application of ratio test. The advantage of the above method is that it gives an explicit formula for the limit $Q(\cdot)$.

Remark 4. Using Theorem 1 one can show that there does not exist a large deviation principle (see [5]) for the convergence of averages in a multitype branching process. The details are similar to the single type case (see [3]).

Proof of Theorem 2. Let $Q^{(1)}$ and $Q^{(2)}$ be any two solutions to

$$Q(f(s)) = \gamma Q(s), ||s|| \leq 1, \quad s \neq 1$$

and

$$Q^{(1)}(0) = 0 = Q^{(2)}(0) \text{ and } \lim_{s \to 0} Q'^{(1)}(s) = P_0 = \lim_{s \to 0} Q'^{(2)}(s).$$

Consider

$$\begin{aligned}
||Q^{(1)}(s) - Q^{(2)}(s)|| &= \gamma^{-1}||Q^{(1)}(f(s)) - Q^{(2)}(f(s))|| \\
&= \gamma^{-n}||Q^{(1)}(f^{(n)}(s)) - Q^{(2)}(f^{(n)}(s))|| \\
&\leq \gamma^{-n}\Big[||Q^{(1)}(f^{(n)}(s)) - Q^{(1)}(0) - f^{(n)}(s)P_0|| \\
&\qquad + ||Q^{(2)}(f^{(n)}(s)) - Q^{(2)}(0) - f^{(n)}(s)P_0||\Big] \\
&= \gamma^{-n}||f^{(n)}(s)|| \Bigg[\frac{||Q^{(1)}(f^{(n)}(s)) - Q^{(1)}(0) - f^{(n)}(s)P_0||}{||f^{(n)}(s)||} \\
&\qquad + \frac{||Q^{(2)}(f^{(n)}(s)) - Q^{(2)}(0) - f^{(n)}(s)P_0||}{||f^{(n)}(s)||}\Bigg]
\end{aligned}$$

Letting $n \to \infty$, we see that for $||s|| \leq 1, s \neq 1$, the first and the second term converge to 0 while $\gamma^{-n}||f^{(n)}(s)|| \to ||Q(s)|| < \infty$ thus completing the proof of Theorem 2.

□

Proof of Theorem 3. We shall only show that

$$\frac{\log f_1^{(n)}(s)}{2^n} \to R_1(s).$$

It is easy to see that

$$f_1^{(1)}(s_1,s_2) = s_1^2 p[1+\delta_1 a(s_2) + \delta_2 b(s_1,s_2)] \tag{7}$$

where

$$\begin{aligned} a(s_2) &= \sum_{j\geq 1} \frac{P_1(2,j)}{\alpha_1} s_2^j & b(s) &= \sum_{i>2}\sum_{j\geq 0} \frac{P_1(i,j)}{\alpha_2} s_1^{i-2} s_2^j \\ \alpha_1 &= \sum_{j\geq 1} P_1(2,j) & \alpha_2 &= \sum_{i>2}\sum_{j\geq 0} P_1(i,j) \\ \delta_1 &= \frac{\alpha_1}{p} & \delta_2 &= \frac{\alpha_2}{p}, \quad p = P_1(2,0) \end{aligned}$$

Note that $a(\cdot)$ and $b(\cdot,\cdot)$ are probability generating functions.

Thus

$$f_1^{(n+1)}(s_1,s_2) = (f_1^{(n)}(s_1,s_2))^2 p[1+\delta_1 a(f_2^{(n)}(s_1,s_2)) + \delta_2 b(f^{(n)}(s_1,s_2))] \tag{8}$$

Define

$$h^{(n)}(s_1,s_2) = (f_1^{(n)}(s_1,s_2))^{\frac{1}{2^n}}.$$

Then

$$h^{(n+1)}(s_1,s_2) = h^{(n)}(s_1,s_2)(L_n(s_1,s_2))^{\frac{1}{2^{n+1}}} \tag{9}$$

where

$$L_n(s_1,s_2) = p\left[1+\delta_1 a(f_2^{(n)}(s_1,s_2)) + \delta_2 b(f^{(n)}(s_1,s_2))\right].$$

Iterating (6) we have

$$h^{(n+1)}(s_1,s_2) = s_1 \prod_{j=1}^{n} (L_j(s_1,s_2))^{\frac{1}{2^{j+1}}}$$

and hence

$$f_1^{(n)}(s_1,s_2) = s_1^{2^n}\left(\prod_{j=1}^{n}(L_j(s_1,s_2))^{2^{-(j+1)}}\right)^{2^n}$$

Now,

$$\frac{\log f_1^{(n)}(s_1,s_2)}{2^n} = \log s_1 + \sum_{j=1}^{n}\frac{1}{2^{j+1}}\log(L_j(s_1,s_2))$$

converges (uniformly in $(s_1,s_2)\in\mathcal{C}_2$)

$$\begin{aligned}(\text{since } |\log L_j(s_1,s_2)| &= |\log p| + \log(1+\delta_1 a(f_1^{(j)}(s_1,s_2)) + \delta_2 b(f^{(j)}(s_1,s_2))))\\ &\leq |\log p| + \log(1+\delta_1+\delta_2) < \infty\end{aligned}$$

We denote the limit $R_1(s_1,s_2)$.

Similar calculations for $f_2^{(n)}$ gives

$$\lim_{n\to\infty}\frac{\log f_2^{(n)}(s)}{2^n} = R_2(s).$$

That R satisfies the functional equation follows easily.

$\square$

Proof of Theorem 4. Consider for fixed $\varepsilon > 0$

$$\begin{aligned}&P_k\left[\left|\frac{\ell\cdot Z_{n+1}}{1\cdot Z_n} - \frac{\ell\cdot(Z_nM)}{1\cdot Z_n}\right| > \varepsilon\right]\\ &= \sum_{j\in\mathcal{A}_2} P(|\ell\cdot Z_{n+1} - \ell\cdot(Z_nM)| > \varepsilon(1\cdot j))P_k(Z_n = j)\end{aligned}$$

Note that

$$\ell\cdot Z_{n+1} = \sum_{k=1}^{2}\ell_k\left(\sum_{r=1}^{1}\sum_{m=1}^{j_r} X_{n,m,r}^{(k)}\right)$$

where $X_{n,m,r}^{(k)}$ is the number of type k offspring coming from the mth parent of type r in the nth generation. Note that for fixed k, r and n, $\{X_{n,m,r}^{(k)} : m = 1, 2 \ldots\}$ are i.i.d. and for $j \in \mathcal{A}_2$

$$\ell \cdot Z_{n+1} - \ell \cdot (jM) = \sum_{k=1}^{2} \ell_k \left(\sum_{r=1}^{2} \left(\sum_{m=1}^{j_r} \tilde{X}_{n,m,r}^{(k)} \right) \right)$$

where $\tilde{X}_{n,m,r}^{(k)} = X_{n,m,r}^{(k)} - m_{rk}$ and $E\left(\tilde{X}_{n,m,r}^{(k)}\right) = 0$ for all n. Thus

$$P\left[\left| \frac{\ell \cdot Z_{n+1}}{1 \cdot Z_n} - \frac{\ell \cdot (Z_n M)}{1 \cdot Z_n} \right| > \varepsilon | Z_n = j \right] = P\left[\left| \sum_{k=1}^{p} \ell_k Y_{n,k} \right| > \varepsilon (1 \cdot j) \right] \triangleq \phi(j, \varepsilon)$$

$$\text{where } Y_{n,k} = \sum_{r=1}^{2} \sum_{m=1}^{j_r} \tilde{X}_{n,m,r}^{(k)}$$

Now, by Markov's inequality, for $\theta > 0$,

$$\begin{aligned} P(\textstyle\sum \ell_k Y_{n,k} > \varepsilon(1 \cdot j)) & \\ &\leq E\left(e^{\theta \sum_{k=1}^{2} \ell_k Y_{n,k}} \right) e^{-\varepsilon\theta(1 \cdot j)} \\ &= \prod_{r=1}^{2} E\left(e^{\theta \sum_{k=1}^{2} \sum_{m=1}^{j_r} \tilde{X}_{n,m,r}^{(k)}} \right) e^{-\varepsilon\theta(1 \cdot j)} \\ &= \prod_{r=1}^{2} (M_r(\theta))^{j_r} e^{-\varepsilon\theta(1 \cdot j)} \\ \text{where } M_r(\theta) &= E\left(e^{\theta \sum_{k=1}^{2} \tilde{X}_{1,1,r}^{(k)}} \right) \\ &= e^{\sum_{r=1}^{2} j_r (\ell n M_r(\theta) - \varepsilon\theta)} \end{aligned}$$

Since $M_r(\theta)$ is the mgf of a mean 0 random variable, there is an appropriate choice of $\theta > 0$ such that

$$e^{\sum_{r=1}^{2} j_r (\ell n Mr(\theta) - \varepsilon\theta)} \leq e^{-(1 \cdot j)\theta C_\varepsilon} \text{ where } C_\varepsilon$$

is a positive constant and hence

$$P(\sum \ell_k Y_{n,k} > \varepsilon(1\cdot j)) \leq C(s^*)^{(1\cdot j)} \text{ for some } 0 < s^* < 1.$$

Using similar calculations for $P(\sum \ell_k, Y_{n,k} < -\varepsilon(1\cdot j))$ we see that

$$\phi(j,\varepsilon) \leq (s^*)^{1\cdot j} \text{ for some } 0 < s^* < 1$$

Now $\phi(j,\varepsilon)\dfrac{P(Z_n = j)}{\gamma^n} \leq C(s^*)^{1\cdot j}\dfrac{P(Z_n = j)}{\gamma^n}$. The left hand side converges to $\phi(j,\varepsilon)q_j$ and the right hand side converges to $(s^*)^{1\cdot j}q_j$ as $n \to \infty$ and $\sum_{j\in\mathcal{A}_2} (s^*_\varepsilon)^{1\cdot j} q_j < \infty$ (by Theorem 1).

The first assertion of Theorem 4 now follows by an application of a generalized version of LDCT to the counting measure space $\mathcal{A}_2$.

For the second part, let k_0 be fixed (to be chosen later). For $n > k_0$,

$$\begin{aligned}
&P_i\left(|\frac{\ell\cdot Z_n}{1\cdot Z_n} - \frac{\ell\cdot v^{(1)}}{1\cdot v^{(1)}}| > \varepsilon\right)\\
&= E_i\left(P\left(|\frac{\ell\cdot Z_n}{1\cdot Z_n} - \frac{\ell\cdot v^{(1)}}{1\cdot v^{(1)}}| > \varepsilon | Z_{n-k_0}\right)\right)\\
&= \sum_{j\in\mathcal{A}_2} P\left(|\frac{\ell\cdot Z_n}{1\cdot Z_n} - \frac{\ell\cdot v^{(1)}}{1\cdot v^{(1)}}| > \varepsilon | Z_{n-k_o} = j\right) P(Z_{n-k_o} = j)
\end{aligned}$$

Consider the event $\left(\dfrac{\ell\cdot Z_n}{1\cdot Z_n} - \dfrac{\ell\cdot v^{(1)}}{1\cdot v^{(1)}}\right) > \varepsilon$ conditioned on $Z_{n-k_0} = j$. By branching property

$$Z_n = \sum_{r=1}^{j_1} Z^{(1)}_{k_o,r} + \sum_{r=1}^{j_2} Z^{(2)}_{k_o,r} \tag{10}$$

where for fixed i, i.i.d., $\{Z^{(i)}_{k_0,r}\}_{r=1}^{\infty}$ are $\mathcal{A}_2$ valued random variables distributed as the population at time k_0 initiated by a particle of type i at time 0. Now,

$$\begin{aligned}
&\frac{\ell\cdot Z_n}{1\cdot Z_n} > \left(\frac{\ell\cdot v^{(1)}}{1\cdot v^{(1)}}+\varepsilon\right)\\
\Leftrightarrow &\ell\cdot Z_n - \ell\cdot(jM^{k_0}) > \left(\frac{\ell\cdot v^{(1)}}{1\cdot v^{(1)}}+\varepsilon\right)(1\cdot Z_n - 1\cdot(jM^{k_0}))\\
&\qquad\qquad + \left(\frac{\ell\cdot v^{(1)}}{1\cdot v^{(1)}}+\varepsilon\right)(1\cdot(jM^{k_0})) - \ell\cdot(jM^{k_0})\\
\Leftrightarrow &\left(\ell - \left(\frac{\ell\cdot v^{(1)}}{1\cdot v^{(1)}}+\varepsilon\right)1\right)\cdot\frac{(Z_n-(jM^{k_0}))}{(u^{(1)}\cdot j)\rho^{k_0}} > \left(\frac{\ell\cdot v^{(1)}}{1\cdot v^{(1)}}+\varepsilon\right)\frac{(1\cdot(jM^{k_0}))-\ell\cdot(jM^{k_0})}{(u^{(1)}\cdot j)\rho^{k_0}}.
\end{aligned}$$

From Frobenius theorem (see [2] lemma 1, page 194) it is known that if $F = \{x=(x_1,x_2)|x_i>0, x\cdot u^{(1)}=1\}$ then

$$\lim_{n\to\infty}\sup_{x\in F} xM^n\rho^{-n} = v^{(1)}.$$

Consequently, there is $k_0<\infty$ such that

$$\sup_{x\in F}||xM^{k_0}\rho^{-k_0} - v||(2||\ell||+\varepsilon) < \varepsilon/2.$$

$$\frac{\ell\cdot Z_n}{1\cdot Z_n} > \frac{\ell\cdot v^{(1)}}{1\cdot v^{(1)}}+\varepsilon$$

$$\text{Then } \left(\ell - \frac{\ell\cdot v^{(1)}}{1\cdot v^{(1)}}+\varepsilon\right)1\cdot\frac{(Z_n - jM^{k_0})}{(u^{(1)}\cdot j)\rho^{k_0}} > \frac{\varepsilon}{2}$$

and hence

$$\begin{aligned}
&P\left(\frac{\ell\cdot Z_n}{1\cdot Z_n} > \frac{\ell\cdot v^{(1)}}{1\cdot v^{(1)}}+\varepsilon\right)\\
&\le e^{-\theta\frac{(u^{(1)}\cdot j)\varepsilon}{2}} e^{j_1\log M_{k_0^1}(\theta)+j_2\log M_{k_0^2}(\theta)}\\
&\text{where } M_{k_0^i}(\theta) = E_i\left(e^{\theta\left(\ell-\frac{\ell\cdot v^{(1)}}{1\cdot v^{(1)}}+\varepsilon\right)\cdot\left(\frac{Z_{k_0}-e_iM^{k_0}}{\rho^{k_0}}\right)}\right).
\end{aligned}$$

Since

$$E_i\left(1\cdot(Z_{k_0}-e_iM^{k_0})\right)=0,$$
$$\lim_{\theta\downarrow 0}\frac{\log M_{k_0^i}(\theta)}{\theta}=0 \text{ for } i=1,2.$$

Thus for any $\varepsilon>0$, there exists $\theta_0>0$ such that

$$\theta_0\frac{u^{(1)}\cdot j\varepsilon}{2}-j_1\log M_{k_0^1}(\theta_0)-j_2\log M_{k_0^2}(\theta)>\theta_0\frac{(u^{(1)}\cdot j)\varepsilon}{4}.$$

This yields the estimate that

$$P\left(\frac{\ell\cdot Z_n}{1\cdot Z_n}-\frac{\ell\cdot v}{1\cdot v}>\varepsilon\ ||Z_{n-k_0}=j\right)\le C(s^{**})^{1\cdot j}$$

for some $0<s^{**}<1$.

Now, using similar calculations for the other side

$$P\left(|\frac{\ell\cdot Z_n}{1\cdot Z_n}-\frac{\ell\cdot v^{(1)}}{1\cdot v^{(1)}}|>\varepsilon|Z_{n-k_0}=j\right)\le C(s^{**})^{1\cdot j}$$

and arguments similar to those in part (i) lead to the desired conclusion.

□

Proof of Theorem 5.

Using central limit theorem and the decomposition as sums of independent random variables (as in the proof of theorem 4) one can see that

$$P\left(|\frac{\ell\cdot Z_{n+1}}{1\cdot Z_n}-\frac{\ell\cdot(Z_nM)}{1\cdot(Z_nM)}|>\varepsilon|Z_n=j\right)\le\frac{C(r,\varepsilon)}{(1\cdot j)^r}$$

where $C(r,\varepsilon)$ is a constant depending only on r and ε.

Thus

$$\frac{1}{\gamma^n}E_1\left(P\left(|\frac{\ell\cdot Z_{n+1}}{1\cdot Z_n}-\frac{\ell\cdot(Z_nM)}{1\cdot Z_n}|>\varepsilon|Z_n\right)\right)\le\frac{C}{\gamma^n}E_1\left(\frac{1}{1\cdot Z_n}\right)^r$$

For any positive random variable X (see [1] for more details) and $0 < r < \infty$

$$\Gamma(r)E(X^{-r}) = 1 \int_0^\infty E(e^{-tX})t^{r-1}dt \tag{13}$$

where

$$\Gamma(r) = \int_0^\infty e^{-x}x^{r-1}dx$$

Applying to $E(1 \cdot Z_n)^{-r}$, we have

$$\Gamma(r)E_i(1 \cdot Z_n)^{-r} = \int_0^\infty f_i^{(n)}(e^{-t}, e^{-t})t^{r-1}dt$$

and hence

$$A_n \equiv \Gamma(r)\frac{(v_1^{(0)}E_1(1 \cdot Z_n)^{-r} + v_2^{(0)}E_2(1 \cdot Z_n)^{-r})}{\gamma^n} \leq C\int_0^\infty \frac{v^{(0)} \cdot f^{(n)}(e^{-t}, e^{-t})t^{r-1}dt}{\gamma^n}.$$

But since $f(s) \geq As$ for s in $\mathcal{C}_2, v^{(0)} \cdot f(s) \geq \gamma v^{(0)} \cdot s$ implying $\dfrac{v^{(0)} \cdot f^{(n)}(s)}{\gamma^n}$ is increasing in n. Now by the monotone convergence theorem

$$\begin{aligned}\lim_{n\to\infty} A_n &= \int_0^\infty v^{(0)} \cdot Q(e^{-t}, e^{-t})t^{r-1}dt \\ &= \int_0^1 v^{(0)} \cdot Q(x,x)K(x)dx \\ \text{where } K(x) &= \frac{|\log x|^{r-1}}{x}\end{aligned}$$

we need to show that $\int_0^1 v^{(0)} \cdot Q(x,x)K(x)dx < \infty$. We start with the following claim.

Claim. *There is an x_0 in $(0,1)$ such that $f_1(x,x) - f_2(x,x)$ is of the same sign for all $x_0 \le x \le 1$.*

Proof. If $m_{11} + m_{12} > m_{21} + m_{22}$ then

$$\lim_{x \to 1} \frac{f_1(x,x) - f_2(x,x)}{1-x} = (m_{21} + m_{22}) - (m_{11} + m_{12}) < 0.$$

Thus there exists x_o such that for all $x \ge x_o \quad f_1(x,x) \le f_2(x,x)$. If $m_{11} + m_{12} = m_{21} + m_{22}$ argue using a higher moment. Let

$$h(x) = \max f_i(x,x) \text{ for } x \text{ in } (x_0, 1).$$

We shall only show that

$$\int_{x_o}^{1} Q_1(x,x) K(x) dx \quad < \infty. \tag{14}$$

From the functional equation and the above claim it follows that $\gamma Q_1(x,x) = Q_1(f(x,x)) \le Q_1(h(x), h(x))$. Let $g(x) = h^{-1}(x)$, then its kth iterate $g_k(x)$ increases to 1 and let $t_k = g_{k-1}(x^*)$, $k \ge 1$ ($x^*$to be chosen later).

Then

$$\int_{x_0}^{1} Q_1(x,x)K(x)dx = \left(\int_{x_0}^{t_0} + \int_{t_0}^{1} \right) Q_1(x,x)K(x)dx$$

$$\text{But } \int_{t_0}^{1} Q_1(x,x)K(x)dx = \sum_{k \ge 0} \int_{t_k}^{t_{k+1}} Q_1(x,x)K(x)dx$$

$$\text{Let } I_k = \int_{t_k}^{t_{k+1}} Q_1(x,x)K(x)dx \le \frac{1}{\gamma} \int_{t_k}^{t_{k+1}} Q_1(h(x),h(x))x(x)dx$$

$$= \int_{t_k}^{t_{k+1}} Q_1(x,x) \frac{K(h^{-1}(x))K(x)dx}{\gamma h(h^{-1}(x))K(x)}$$

$$\text{But } \frac{K(h^{-1}(x))}{\gamma h'(h^{-1}(x))K(x)} \to \frac{1}{\gamma (h'(1))^r} = \beta < 1 \text{ by A6 .}$$

Hence there exists x_0 such that for $x_0 \le x \le 1$

$$\frac{K(h^{-1}(x))}{\gamma h^1(h^{-1}(x))K(x)} < (\beta+\varepsilon) < 1. \tag{15}$$

Now choosing $x^* = x_0$ we have

$$I_k \le (\beta+\varepsilon) \int_{t_{k-1}}^{t_k} Q_1(x,x)K(x)dx$$

$$I_k \le (\beta+\varepsilon) I_{k-1}$$
$$\le (\beta+\varepsilon)^k I_0$$

Thus

$$\sum_{k\ge 1} \int_{t_k}^{t_{k+1}} Q_1(x,x)K(x)dx \le \frac{I_0}{(1-(\beta+\varepsilon))} < \infty$$

completing the proof of (14). The proof of the first part of the main theorem is completed by using the same arguments as in Theorem 4. The proof of the second part is similar.

□

Proof of Theorem 6.

Since $E\left(e^{\theta_0(1\cdot Z_1)}\right) < \infty$, given $\varepsilon > 0$ there exists constants $C_1(\varepsilon), s_\varepsilon, C_2(\varepsilon), \tilde{s}_\varepsilon$ such that

$$P_1\left[\left|\frac{\ell\cdot Z_{n+1}}{1\cdot Z_n} - \frac{\ell\cdot(Z_nM)}{1\cdot Z_n}\right| > \varepsilon\right] \le C_1(\varepsilon)\sum_{j\in\mathcal{A}_2} s_\varepsilon^{1\cdot j} P_{e_1}(Z_n=j)$$

$$= C_1(\varepsilon) f_1^{(n)}(s_\varepsilon 1)$$

and

$$P_1\left[\left|\frac{\ell\cdot Z_n}{1\cdot Z_n} - \frac{\ell\cdot v}{1\cdot v}\right| > \varepsilon\right] \le C_2(\varepsilon) f_1^{(n)}(\tilde{s}_\varepsilon\cdot 1)$$

□

The result follows from Theorem 3.

Proof of Theorem 7. Let $\phi_{n,1}(\theta) = E_1(e^{\theta W_n})$. Note that $\phi_{n,1}(\theta)$ is monotone in θ. Let $\{\theta_n \colon n \geq 1\}$ be such that

$$E\left(e^{\theta_n W_n}\right) = E\left(e^{\theta_1 W_1}\right) = k < \infty.$$

We start the proof with the following lemmas.

□

Lemma 3. *$\{\theta_n \colon n \geq 1\}$ is a monotone decreasing sequence.*

Proof. For $n \geq 1$

$$\begin{aligned} k = E\left(e^{\theta_n W_n}\right) &= E\left(E(e^{\theta_{n+1} W_{n+1}} | \mathcal{F}_n)\right) \\ &\geq E\left(e^{\theta_{n+1} W_n}\right) \end{aligned}$$

where the last inequality follows from conditional Jensen's inequality (see [6]). Thus

$$\theta_n \geq \theta_{n+1} \qquad \text{for all } n \geq 1$$

completing the proof of the lemma.

From the claim it follows that $\lim_{n \to \infty} \theta_n \equiv \theta_0^*$ exists. Thus it is enough to show that $\theta_0^* > 0$.

The estimate (16) needed in the proof of our main theorem is the content of our next lemma.

□

Lemma 4. *Let*

$$\tilde{\Psi}_r(\theta) = E_r\left(e^{\theta(u^{(1)} \cdot Z_1 - u_r^{(1)} \rho)}\right)$$

Then

$$\tilde{\Psi}_r(\theta) \leq 1 + C\theta^2 \quad \text{for } r = 1, 2 \tag{16}$$

Proof. Let $X \equiv u^{(1)} \cdot Z_1 - u_r^{(1)}\rho$. Then $E_r(X) = 0$ and

$$|\frac{\tilde{\Psi}_r(\theta) - 1}{\theta^2}| = |\frac{E_r(e^{\theta X} - x - 1)}{\theta^2}|$$

$$\leq |E_r\left(\frac{e^{\theta X} - x - 1}{\theta^2 X^2} X^2 \colon |\theta^2 X^2| \leq 1\right)| + |E_r\left(\frac{C^{\theta X} - x - 1}{\theta^2 X^2} X^2 \colon |\theta^2 X^2| > 1\right)|$$

$$\leq C_1 E X^2 + C_2 = C < \infty \text{ where } C_1 = \sup_{|u| \leq 1} \frac{|e^u - u - 1|}{u^2}$$

$$\text{and } C_2 = E\left[\left(e^{\theta_0/2u^{(1)} \cdot Z_1} + 1 + u^{(1)} \cdot Z_1 + u_r^{(1)}\rho\right)\left(u^{(1)} \cdot Z_1 + u_r^{(1)}\rho)\right)^2\right]$$

□

We now return to the proof of the main theorem.

Consider

$$E\left(e^{\theta_{n+1} W_{n+1}}\right) = E\left(e^{\theta_{n+1} W_n} E\left(e^{\theta_{n+1}(W_{n+1} - W_n)} \mid \mathcal{F}_n\right)\right)$$

Note that

$$W_{n+1} - W_n = \rho^{-(n+1)} \sum_{r=1}^{2} \sum_{j=1}^{Z_n^{(r)}} (u^{(1)} \cdot \xi_{n,j,r} - u_r^{(1)}\rho)$$

where

$$\xi_{n,j,r} = (\xi_{n,j,r}^{(1)}, \xi_{n,j,r}^{(2)})$$

is the number of offspring of a jth type r parent in the nth generation.

Thus

$$
\begin{aligned}
&E\left(e^{\theta_{n+1}(W_{n+1}-W_n)} \mid \mathcal{F}_n\right)\\
&= E\left(exp\left\{\frac{\theta_{n+1}}{\rho^{n+1}}\sum_{r=1}^{2}\sum_{j=1}^{Z_n^{(r)}}\left(u^{(1)}\cdot\xi_{n,j,r}-u_r^{(1)}\rho\right)\right\}\right)\\
&= \prod_{r=1}^{2}\left(\tilde{\Psi}_r\left(\frac{\theta_{n+1}}{\rho^{n+1}}\right)\right)^{Z_n^{(r)}}\\
&\leq e^{\sum_{r=1}^{2} Z_n^{(r)}||\log\tilde{\Psi}.\left(\frac{\theta_{n+1}}{\rho^{n+1}}\right)||}\\
&\leq e^{CW_n\left(||\log\tilde{\Psi}.(\theta_{n+1}/\rho^{n+1})||\theta_{n+1}/(\theta_{n+1}/\rho^{n+1}\right)}.
\end{aligned}
$$

Hence

$$
E\left(e^{\theta_{n+1}W_{n+1}}\right) \leq E\left(e^{\theta_{n+1}W_n\left(1+C\frac{||\log\tilde{\Psi}.(\theta_{n+1}/\rho^{n+1})}{\theta_{n+1}/\rho^{n+1}}\right)}\right)
$$

which implies that

$$
\theta_{n+1} \geq \frac{\theta_n}{\left(1+\frac{C||\log\tilde{\Psi}.(\theta_{n+1}/\rho^{n+1})||}{\theta_{n+1}/\rho^{n+1}}\right)}.
$$

Iterating the above we have

$$
\theta_{n+1} \geq \theta_1\left(\prod_{j=1}^{n+1}\left(1+\frac{C||\log\tilde{\Psi}.(\theta_j/\rho^j)}{\theta_j/\rho^j}\right)\right)^{-1}.
$$

Letting $n \to \infty$ we have

$$
\theta_0^* \geq \theta_1\left(\prod_{j=1}^{\infty}\left(1+\frac{C||\log\tilde{\Psi}.(\theta_j/\rho^j)||}{\theta_j/\rho^j}\right)\right)^{-1} > 0
$$

where the positivity of the last term follows from lemma 4.

Remark 5. *Similar results for single type branching processes were used by Biggins and Shanbhag (see [4]) for studying infinite divisibility problems in single type branching processes.*

Proof of Theorem 8.

Consider

$$\begin{aligned} W - W_n &= \lim_{k\to\infty} (W_{n+k} - W_n) \\ &= \lim_{k\to\infty} \left(\frac{1}{\rho^n} \sum_{r=1}^{2} \sum_{j=1}^{Z_n^{(r)}} \left(\frac{u^{(1)} \cdot Z_{k,j}}{\rho^k} - u_r^{(1)} \right) \right) \\ &= \lim_{k\to\infty} \frac{1}{\rho^n} \sum_{r=1}^{2} \sum_{j=1}^{Z_n^{(r)}} \left(W_{k,j}^{(r)} - u_r^{(1)} \right) \\ &= \frac{1}{\rho^n} \sum_{r=1}^{2} \sum_{j=1}^{Z_n^{(r)}} \left(W_j^{(r)} - u_r^{(1)} \right) \end{aligned}$$

Note that for fixed $r, (W_j^{(r)}, j = 1, 2 \ldots, Z_n^{(r)})$ are i.i.d. and if $r_1 \neq r_2, W_j^{(r_1)}$ and $W_j^{(r_2)}$ are independent. Consider

$$P(W - W_n > \varepsilon | \mathcal{F}_n) = P\left(\sum_{r=1}^{2} \sum_{j=1}^{k_r} Y_j^{(r)} > \varepsilon \right) \triangleq \Phi(k, \varepsilon)$$

where $Y_j^{(r)} = W_j^{(r)} - u_r^{(1)}$, $\;k = (k_1, k_2)$. Then $P_1(W - W_n > \varepsilon) = E_1 \Phi(Z_n, \rho^n \varepsilon)$.

$$\begin{aligned}
\text{Consider } \Phi(k,\varepsilon) &= P\left(\frac{1}{\sqrt{u^{(1)}\cdot k}} \sum_{r=1}^{2} \sum_{j=1}^{k_r} Y_j^{(r)} > \frac{\varepsilon}{\sqrt{u^{(1)}\cdot k}}\right) \\
&\le E\left(e^{\frac{\theta}{\sqrt{u^{(1)}\cdot k}} \sum_{r=1}^{2}\sum_{j=1}^{k_r} Y_j^{(r)}}\right) e^{-\theta\varepsilon/\sqrt{u^{(1)}\cdot k}} \\
&= \prod_{r=1}^{2} E_r\left(e^{\frac{\theta}{\sqrt{u^{(1)}\cdot k}} \sum_{j=1}^{k_r} Y_j^{(r)}}\right) e^{-\theta\varepsilon/\sqrt{u^{(1)}\cdot k}} \\
&\triangleq \left(\prod_{r=1}^{2}\left(\Psi_r\left(\frac{\theta}{\sqrt{u^{(1)}\cdot k}}\right) e^{\frac{\theta u_r^{(1)}}{\sqrt{u^{(1)}\cdot k}}}\right)^{k_r}\right) e^{-\frac{\theta\varepsilon}{\sqrt{u^{(1)}\cdot k}}}
\end{aligned}$$

Note that

$$\left(\Psi_r\left(\frac{\theta}{\sqrt{u^{(1)}\cdot k}} e^{-\frac{\theta u_r^{(1)}}{\sqrt{u^{(1)}\cdot k}}}\right)^{k_r}\right) = \left(1 + \frac{1}{k_r}\frac{\left(\Psi_r\left(\frac{\theta}{\sqrt{u^{(1)}\cdot k}}\right) e^{-\frac{\theta u_r^{(1)}}{\sqrt{u^{(1)}\cdot k}}} - 1\right)}{(\theta^2/k_r)}\theta^2\right)^{k_r}$$

and observe that

$$u_r^{(1)} \sup_{0\le\theta\le\theta_2} \frac{|\Psi_r\left(\frac{\theta}{\sqrt{u^{(1)}\cdot k}}\right) e^{-\theta u_r^{(1)}/\sqrt{u^{(1)}\cdot k}} - 1|}{\left(\theta^2/u_r^{(1)}k_r\right)} \text{ is bounded above by}$$

$$u_r^{(1)} \sup_{0\le\theta\le\theta_2} \frac{|\Psi_r\left(\frac{\theta}{\sqrt{u^{(1)}\cdot k}}\right) e^{-\theta u_r^{(1)}/\sqrt{u^{(1)}\cdot k}} - 1|}{\left(\theta^2/u^{(1)}\cdot k\right)} \equiv C_r < \infty$$

where $\theta_2 = min(\theta_0, 1)$. Thus $\left(\Psi_r\left(\frac{\theta}{\sqrt{u^{(1)}\cdot k}}\right)e^{-\theta u_r^{(1)}/\sqrt{u^{(1)}\cdot k}}\right)^{k_r} \le e^{C_r}$

and hence $\Phi(k,\varepsilon) \le Ce^{\frac{-\varepsilon\theta}{\sqrt{u^{(1)}\cdot k}}}$.

$$\text{Thus} \quad P_1(W - W_n > \varepsilon) = E_1(\Phi(Z_n, \rho^n\varepsilon)) \le CE_1\left(e^{-\rho^n\varepsilon\theta/\sqrt{u^{(1)}\cdot Z_n}}\right)$$

$$= CE_1\left(e^{-\rho^{n/2}\varepsilon\theta/\sqrt{W_n}}\right).$$

$$\text{Now} \quad E_1\left(e^{-\gamma_1/\sqrt{W_n}}\right) = \gamma_1\int_0^\infty e^{-\gamma_1 u}P_1\left(\frac{1}{\sqrt{W_n}} \le u\right)du$$

$$= \gamma_1\int_0^\infty e^{-\gamma_1 u}P_1\left(W_n \ge \frac{1}{u^2}\right)du$$

$$\le \gamma_1 C_1\int_0^\infty e^{-\gamma_1 u}e^{-\theta/u^2}du = C_1\int_0^\infty e^{-t}e^{\frac{-\theta\gamma_1^2}{t^2}}dt \le 2e^{-\gamma_1^{2/3}}$$

Thus

$$P_1(W - W_n > \varepsilon) \le 2C\, e^{\rho^{n/3}\varepsilon^{2/3}\theta^{2/3}}.$$

Similar arguments hold for $P_1(W_n - W > \varepsilon)$.

Proof of Theorem 9.

Note that $P\left(|\frac{\ell\cdot Z_{n+1}}{1\cdot Z_n} - \frac{\ell\cdot(Z_nM)}{1\cdot Z_n}| > \varepsilon \mid W \ge a\right)$ is the same as

$$P\left(|\frac{\ell\cdot Z_{n+1}}{1\cdot Z_n} - \frac{\ell\cdot(Z_nM)}{1\cdot Z_n}| > \varepsilon, W \ge a\right)\cdot\frac{1}{P(W \ge a)}.$$

The above is the same as

$$p_a\left[P\left(|\frac{\ell\cdot Z_{n+1}}{1\cdot Z_n}-\frac{\ell\cdot(Z_nM)}{1\cdot Z_n}|>a, W_n\geq a\gamma_2, W\geq a\right)\right.$$
$$\left.+P\left(|\frac{\ell\cdot Z_{n+1}}{1\cdot Z_n}-\frac{\ell\cdot(Z_nM)}{1\cdot Z_n}|>a, W_n\leq a\gamma_2, W\geq a\right)\right]$$

where $0<\gamma_2<1$, and $p_a=P(W\geq a)$.

The second term inside the parenthesis is bounded above by

$$P(W-W_n\geq a(1-\gamma_2))\leq C_4e^{-\gamma_2(a(1-\gamma_2))^{2/3}(\rho^{1/3})^n} \text{ (by Theorem 6)}$$

As for the first term, note that it is bounded above by $C_5e^{-C_\epsilon a\gamma_2\rho^n}$ (by using Chernoff type bounds). Combining the estimates we have that

$$P\left(|\frac{\ell\cdot Z_{n+1}}{1\cdot Z_n}-\frac{\ell\cdot(Z_nM)}{1\cdot Z_n}|>\varepsilon|W\geq a\right)\leq C_4e^{-\gamma_2(a(1-\gamma_2))^{2/3}(\rho^{1/3})^n}+C_5e^{-C_\epsilon a\gamma_2^n\rho^n}.$$

The second part is similar.

□

5. Open Questions.

There are several unresolved questions that arise from our work. We list a few of them:

1) We showed that if A is indecomposable and positively regular then the Perron's root of A determines the growth rate of $f_n(\cdot)$. What would be the decay rate of $f_n(\cdot)$ if A were decomposable?
2) We showed that if A is a zero matrix and satisfies the conditions of Theorem 3 then $f_n(\cdot)$ grows like $2^nR(\cdot)$. Suppose that $P_1(2,0)P_1(0,2), P_2(2,0)$ and $P_2(0,2)$ are positive; then set

$$C=\begin{pmatrix}P_1(2,0) & P_1(0,2)\\ P_2(2,0) & P_2(0,2)\end{pmatrix}$$

and one can write

$$f(s) = Cs^2 + g(s)$$

Note that iterates of $f(\cdot)$ are non-linear in nature. One would like to conjecture that in this case the growth rate of $f^{(n)}(\cdot)$ should be determined by a parameter which relates Perron's root of C and the mimimum family size of 2. Is such a conjecture true, if so what is that parameter?

3) Is it possible to prove Theorem 4 without any conditions other than the finiteness of the mean matrix? A possible approach to this problem is to approximate the offspring distribution function by distributions with finite support.

References.

(1) Athreya, K. B. (1994): *Large deviation rates for Branching Process - I, the single type case*, Annals of Applied Probability (to appear).

(2) Athreya, K. B. and Ney, P. E. (1972): *Branching processes*, Springer-Verlag, New York.

(3) Athreya, K. B. and Vidyashankar, A. N. (1993): *Large deviation results for branching processes, Stochastic Processes*, a Festschrift in honour of Gopinth Kallianpur, Springer-Verlag.

(4) Biggins, J. D. and Shanbhag, D. N. (1981): *Some divisbility problems in branching processes*, Math. Proc. Camb. Phil. Soc., **90**, 321-330.

(5) Deuschel, J. D. and Stroock, D. W. (1989): *Large Deviations*, Academic Press, New York.

(6) Durett (1991): *Probability: Theory and Examples*, Wadsworth and Brooks/Cole Advanced Books Software, Pacific Grove, California.

(7) Royden, H. L. (1987): *Real Analysis*, MacMillan Publishing Company, New York, Third Edition.

(8) Vidyashankar, A. N. (1994): *Large deviations for the tail behavior of W in*

a multitype branching process, preprint, Department of Mathematics, Iowa State University, Ames, IA.

LARGE DEVIATIONS FOR THE TAIL BEHAVIOR OF W IN A MULTITYPE BRANCHING PROCESS

A PAPER SUBMITTED TO JOURNAL OF THEORETICAL PROBABILITY
ANAND N.VIDYASHANKAR

Departments of Mathematics and Statistics
Iowa State University
Ames, IA 50011

Abstract.

This paper is concerned with the tail behavior of the random variable W that comes up in a multi-type branching process. In particular, we focus on two specific questions: First, we consider the rate of decay of $P(W \leq x)$ as x approaches 0. We do this by studying the density w of W near origin. Second, we investigate the rate of decay of $P(W > x)$ as x approaches ∞ when the branching process has finite support. For this reason we develop multi-type versions of Harris functions and use that to extract the exponential rate of decrease.

1. Introduction.

Let $Z_n = (Z_n^{(1)}, Z_n^{(2)})$ denote a multi-type branching process with finite mean matrix M. Assume that M is irreducible and positively regular with maximal eigenvalue ρ. Let $u^{(1)}$ and $v^{(1)}$ denote the right and left eigenvectors associated with ρ. We shall further assume that $\rho > 1$.

It is known that $W_n = \frac{u^{(1)} \cdot Z_n}{\rho^n}$ is a martingale sequence and hence converges with probability 1 to a non-negative random variable W. Suppose we assume that the extinction probability vector $q = 0$. Then it follows that $P(W \leq x)$ converges to 0 as x decreases to 0. One also knows that $P(W \geq x)$ converges to 0 as $x \to \infty$. The aim of this paper is to study the rates of convergence to zero.

The approach for studying these two questions is based on the recent work of Biggins and Bingham (see [4]). However, the tools required for using this approach have only recently been developed in Athreya and Vidyashankar (see [1]). For studying the decay rates for the right tail we assume that the branching process has finite support.

The paper is organized as follows: section 2 considers the decay rates for the right tail while in section 3 we consider the decay rates for the left tail.

2. Decay rates for the right tail of W.

Set $C_2 = [0,1] \times [0,1]$. Let $f\colon C_2 \to C_2$ be defined by

$$f(s_1, s_2) = (f_1, (s_1 s_2), f_2(s_1, s_2))$$

where, for each k,

$$f_k(s_1, s_2) = \sum_{i \geq 0} \sum_{j \geq 0} P_k(i,j) s_1^i s_2^j \text{ and}$$

$P_k(\cdot,\cdot)$ is a probability distribution on the two dimensional integer lattice. Set

$$e_1 = (1,0) \text{ and } e_2 = (0,1).$$

Our first theorem is a generalization of a result of T. E. Harris (see [7]) to two-type branching processes.

Theorem 1. *Assume that $f(0,0) = 0$, hold $P_i(e_j) = 0$ for $i,j = 1,2$ and $P_1(i,j) > 0$ and $P_2(i,j) > 0$ hold for $i = 1, \ldots, d_1 (< \infty), j = 1, 2, \ldots, d_2 (< \infty)$. Let $\sum_{j=1}^{d_2} \sum_{i=1}^{d_2} P_k(i,j) = 1$ for $k = 1,2$. Let $\eta > 1$ be such that*

$$\rho^\eta = g.$$

where $g = d_1 + d_2$. Then for $x > 0$, $\log E_i e^{xW} = x^\eta H_i(x) + H_{0,i}(x)$ for $i = 1,2$ where $H_i(x)$ has the following properties:

(i) *$H_i(x)$ is positive, convex ,$H_i'(0+) = 0, H_i'(x-) \to \infty$ as $x \to \infty$*

(ii) *$H_i(\cdot)$ is real analytic on $(0,\infty)$*

(iii) *$H_i(\cdot)$ is multiplicatively periodic with period ρ*

(iv) *$H_i(\cdot)$ is strictly convex on $(0,\infty)$.*

(v) *$H_{0,i}(x) = O\left(\frac{1}{x^\eta}\right)$ as $x \to \infty$.*

Proof. We shall prove the theorem only for $i = 1$. Let $s \in [1,\infty) \times [1,\infty)$. Then

$$\begin{aligned} f_1(s) &= \sum_{j_1=1}^{d_1} \sum_{j_2=1}^{d_2} P_1(j_1,j_2) s_1^{j_1} s_2^{j_2} \\ &= P_1(d_1,d_2) s_1^{d_1} s_2^{d_2} \sum_{j_1=1}^{d_1} \sum_{j_2=1}^{d_2} \frac{P_1(j_1,j_2)}{P_1(d_1,d_2)} s_1^{j_1-d_1} s_2^{j_2-d_2} \end{aligned}$$

Define

$$\begin{aligned} R_1(s) &= \sum_{j_1=1}^{d_1} \sum_{j_2=1}^{d_2} \frac{P_1(d_1-j_1,d_2-j_2)}{P_1(d_1,d_2)} s_1^{-j_1} s_2^{-j_2} \\ \text{and } R_2(s) &= \sum_{j_1=1}^{d_1} \sum_{j_2=1}^{d_2} \frac{P_2(d_1-j_1,d_2-j_2)}{P_2(d_1,d_2)} s_1^{-j_1} s_2^{-j_2}. \end{aligned}$$

Thus

$f_1(s) = p_1 s^d R_1(s)$ and $f_2(s) = p_2 s^d R_2(s)$ where $p_1 = P_1(d_1,d_2)$ and $p_2 = P_2(d_1,d_2)$

Note that

$$\begin{aligned} (f(s))^d = p^d (s^d)^g (R(s))^d \text{ where } p = (p_1,p_2), \\ R(s) = (R_1(s), R_2(s)) \text{ and } d = (d_1,d_2). \end{aligned}$$

Iterating, we have for $n \geq 2$,

$$f_1(f^{(n)}(s)) = p_1 p^{d(1+g+\cdots+g^{n-1})} s^{dg^n} \prod_{k=1}^{n} (R(f^{(k-1)}(s)))^{dg^{n-k}}$$

Taking logs and dividing by g^n, we get

$$\begin{aligned}
\frac{1}{d_1 g^n}\log f_1(f^{(n)}(s)) &= \frac{\log p_1}{d_1 g^n} + \frac{(1-g^{-n})}{d_1(g-1)}\left[\log p_1^{d_1} + \log p_2^{d_2}\right] + \frac{1}{d_1}\log s_1^{d_1} + \frac{1}{d_1}\log s_2^{d_2} \\
&\quad + \sum_{k=1}^{n} g^{-k}\log R_1(rf^{(k-1)}(s)) + \sum_{k=1}^{n} g^{-k}\frac{d_2}{d_1}\log R_2(f^{(k-1)}(s)) \\
&= \frac{\log \rho_1}{d_1 g^n} + \left(\frac{1-g^{-n}}{g-1}\right)\log p_1 + \frac{(1-g^{-n})}{(g-1)}\log p_2\left(\frac{d_2}{d_1}\right) \\
&\quad + \log s_1 + \frac{d_2}{d_1}\log s_2 + \sum_{k=1}^{n} g^{-k}\log R_1(f^{(k-1)}(s)) \\
&\quad + \frac{d_2}{d_1}\sum_{k=1}^{n} g^{-k}\log R_2(f^{(k-1)}(s)).
\end{aligned}$$

Letting $n \to \infty$, the right hand side converges to

$$\left(\frac{\log p_1}{g-1}\right) + \log s_1 + \sum_{k\geq 1} g^{-k}\log R_1(f^{(k-1)}(s)) + A(s_1, s_2)$$

where $A(s_1, s_2) = \frac{d_2}{d_1}\left[\frac{\log p_2}{g-1} + \log s_2 + \sum_{k\geq 1} g^{-k}\log R_2(f^{(k-1)}(s))\right]$.
Let $U(s) = U_1(s) + U_2(s)$ where

$$U_1(s) = \frac{\log p_1}{g-1} + \log s_1 + \sum_{k\geq 1} g^{-k}\log R_1(f^{(k-1)}(s)) \text{ and}$$

$$U_2(s) = \frac{d_2}{d_1}\left[\frac{\log p_2}{g-1} + \log s_2 + \sum_{k\geq 1} g^{-k}\log R_2(f^{(k-1)}(s))\right].$$

For $x > 0$, we then have $U(\Psi(x)) = U_1(\Psi(x)) + U_2(\Psi(x))$ where $\Psi(x) = (\Psi_1(x), \Psi_2(x))$.

Define for $k = 1, 2, T_k(U(\Psi(x))) = U_k(\Psi(x))$, and set $L_1(x) = T_1(U(\Psi(x)))$. We shall show that $L_1(\rho x) = g L_1(x)$.

Towards this, note that

$$U(f(\Psi(x))) = \lim_{n\to\infty} \frac{1}{d_1 g^n} \log f_1(f^{(n+1)}(\Psi(x)))$$

$$= g \lim_{n\to\infty} \frac{1}{d_1 g^{n+1}} \log f_1(f^{(n+1)}(\Psi(x)))$$

$$= gU(\Psi(x)) = gU_1(\Psi(x)) + gU_2(\Psi(x)).$$

The above calculations imply that

$$L_1(x\rho) = T_1(U(\Psi(\rho x))) = T_1(U(f(\Psi(x)))) = gL_1(x). \tag{1}$$

Define $H_1(x) = x^{-\eta} L_1(x)$. From (1) and the definition of η it follows that H_1 is multiplicatively periodic with period ρ.

Let

$$|H_{0,1}(x)| = |\frac{\log \Psi_1(x)}{x^\eta} - H_1(x)|$$
$$= |x^{-\eta} \log \Psi_1(x) - x^{-\eta} \left(\frac{\log p_1}{g-1} + \log \Psi_1(x) + \sum_{k\geq 1} g^{-k} \log R_1(f^{(k-1)}(\Psi(x))) \right) |$$
$$\leq x^{-\eta} (\frac{|\log p_1|}{(g-1)} + \sum_{k\geq 1} g^{-k} \log R_1(f^{(k-1)}(\Psi(x)))).$$

Whence

$$|H_{0,1}(x)| = 0\left(\frac{1}{x^\eta}\right) \quad \text{as} \quad x \to \infty \text{ (since } \sum_{k\geq 1} g^{-k} \log R_1(f^{(k-1)}(\Psi(x))) \text{ is bounded}$$

as $x \to \infty$). Thus

$$\frac{E_1(e^{xW})}{x^\eta} = H_1(x) + H_{0,1}(x) \text{ for } x > 0.$$

As for properties of $H_i(\cdot)$, it is enough to establish (i), (ii), and (iv) for $L_1(x)$. Positivity and convexity of L_1 are obvious. Consider

$$\frac{L_1(\delta\rho^{-n})}{\delta\rho^{-n}} = \frac{g^{-n}L_1(\delta)}{\delta\rho^{-n}} = \frac{1}{\delta}(pg^{-1})^n L_1(d) \to 0$$

as $n \to \infty$, giving (i). Similarly

$$\begin{aligned} L_1'(\delta\rho^{n+1}) &\geq \frac{L_1(\delta\rho^{n+1}) - L_1(\delta\rho^n)}{\delta\rho^{n+1} - \delta\rho^n} \\ &= \frac{(g^{n+1} - g^n)}{\rho^n}\frac{L_1(\delta)}{\delta(\rho-1)} \\ &= (g\rho^{-1})^n\left(\frac{g-1}{\rho-1}\right)\frac{L_1(\delta)}{\delta} \\ &\to \infty \text{ as } n \to \infty. \end{aligned}$$

We now turn our attention to real analyticity. Note that

$$L_1(z) = \frac{\log p_1}{g-1} + \log\Psi_1(z) + \sum_{k\geq 1} g^{-k}\log R_1(f^{(k-1)}(z)).$$

We need to show that

$$\{z\colon L_1(z) \text{ is analytic }\} \supset (0,\infty).$$

We'll show that

$$\{z\colon U(\Psi(z)) \text{ is analytic }\} \supset (0,\infty).$$

Define

$$\mathcal{B} = \bigcap_{n\geq 1}\{z \in \mathbb{C}\colon |f_1^{(n)}(z)| > (|f_1^{(n-1)}(z)\ |^\delta \wedge \ |\ f_2^{(n-1)}(z)\ |^\delta\}$$

where $\delta \in (1,\rho)$.

Claim. $\mathcal{B} \supset (1,\infty)$.

Proof. First observe that for all $y_1 \in (1,\infty)$ and $y_2 \in (1,\infty)$

$$f_1(y_1,y_2) \geq (y_1 \wedge y_2)^\delta > 1$$

$$\text{and} \quad f_2(y_1,y_2) \geq (y_1 \wedge y_2)^\delta > 1$$

Thus

$$\begin{aligned} f_1^{(n)}(y,y) &= f_1(f_1^{(n-1)}(y,y), f_2^{(n-1)}(y,y)) \\ &\geq (f_1^{(n-1)}(y,y) \wedge f_2^{(n-1)}(y,y))^{\delta} \end{aligned}$$

concluding the proof of the claim.

Also note that $\dfrac{\log f_1(f^{(n)}(z))}{dg^n}$ converges uniformly on $\mathcal{B}$ as $|z| \to \infty$ to $U(z)$. Since $\Psi(z)$ is analytic in z and $\Psi_i(x) \in (1, \infty)$ for $x > 0$, it follows that U, and hence L_1, is real analytic on $(1, \infty)$. The final part follows from the fact that an analytic function which is linear somewhere, is linear everywhere.

Our next theorem is a large deviation theorem for $P_i(W \geq x)$ as $x \to \infty$.

Theorem 2. *Assume A1 – A3, and A8 hold. Then as* $x \to \infty$

$$-\log P_i(W \geq x) = x^{\eta/(\eta-1)} H_i^+(x) + O\left(x^{\eta/(\eta-1)}\right)$$

where η *is as in Theorem 1 and*

$$H_i^+(x) \equiv x^{-\eta/(\eta-1)} \sup_{\theta} [\theta x - \theta^{\eta} H_i(\theta)\}$$

is real analytic on $[0, \infty)$, *strictly convex, strictly increasing, and is multiplicatively periodic with period* $\rho^{\eta-1}$.

Proof.

From Theorem 1,

$$\lim_{n\to\infty} \frac{\log E_i(e^{x\rho^n W})}{\rho^{n\eta}} = x^{\eta} H_i(x) = G_i(x)$$

Hence using Corollary 1 of [4], we have

$$\lim_{n\to\infty} -\frac{\log P_i(W \geq \rho^{n(\eta-1)} x)}{\rho^{n\eta}} = G_i^*(x)$$

where $G_i^*(x) \triangleq \sup_\theta [\theta x - G_i(\theta)]$. Further, by the differentiability and convexity of $G_i(\cdot)$, there exists functions $T_i(\cdot)$ such that

$$G_i(x) = \int_0^x T_i(t)dt$$

Note that $T_i(\cdot)$ is real analytic and strictly increasing. Let $T_i^{-1}(\cdot)$ be the inverse of $T_i(\cdot)$. Then $T_i(\cdot)$ is real analytic and strictly increasing. From Young's inequality, (see [6]) the Fenchel dual is given by

$$G_i^*(x) = \int_0^x T_i^{-1}(t)dt.$$

Hence $G_i^*(x)$ is real analytic, strictly convex and strictly increasing on $(0, \infty)$.

Finally, if $H_i^+(x) = x^{-\eta/\eta-1} G_i^*(x)$ then it is easy to see that

(i) H_i^+ is real analytic, strictly convex, $H_i^+(x) > 0$ for $x > 0$

(ii) H_i^+ is multiplicatively periodic with period $\rho^{\eta-1}$.

To conclude the proof, let

$$H_{n,i}^+(x) = -\frac{\log P_i(W \geq \rho^{n(\eta-1)}x)}{\rho^{n\eta} x^{\eta/\eta-1}}$$

then $x^{\eta/\eta-1} H_{n,i}^+(x) \to G_i^*(x)$ as $n \to \infty$.

By monotonicity of $H_{n,i}^+(x)$ and the continuity of the limit, the above convergence is uniform over compact subintervals, in particular over $[1, \rho^{\eta-1}] = I$. This implies that

$$\sup_{x \in I} |H_i^+(x) - H_{n,i}^+(x)| \to 0 \qquad \text{as } n \to \infty.$$

As H_i^+ is multiplicatively periodic with period $\rho^{\eta-1}$

$$\sup_{x \in \rho^{n(\eta-1)} I} |H_i^+(x) + \log P_i(W \geq x) x^{-\eta/\eta-1}| = \sup_{x \in I} |H_i^+(x) - H_{n,i}^+(x)|$$

so that

$$\sup_{x \geq \rho^{N(\eta-1)}I} |H_i^+(x) + x^{-\eta/\eta-1} \log P_i(W \geq x)|$$
$$= \sup_{n \geq N} \sup_{x \in I} |H_i^+(x) - H_{n,i}^+(x)| \to 0 \text{ as } n \to \infty.$$

□

3. Decay rates for the left tail of W.

In this section we do not assume the branching process to have finite support. We shall now consider the large deviation of $P_i(W \leq x)$ as $x \to 0$ in the Schröder case. First recall the following theorem from [1].

Theorem 3. *Assume that the matrix* $A = \begin{bmatrix} P_1(1,0) & P_1(0,1) \\ P_2(1,0) & P_2(0,1) \end{bmatrix}$ *is positively regular and irreducible. Let γ be the Perron's root of A. Assume further that $f(0,0) = 0$. Then there exists a map $Q\colon C_2 - 1 \to R_+^2$ such that*

$$\frac{f_n(s)}{\gamma^n} \to Q(s)$$

and Q satisfies the functional equation

$$Q(f(s)) = \gamma Q(s)$$
$$Q(0) = 0$$
$$\text{and } \lim_{s \uparrow 1} Q(s) = \infty$$

Our next theorem gives bounds on the density of W near the origin. These results generalize the results of Dubuc (see [5]) to multitype branching processes.

Theorem 4. *Under the conditions of Theorem 3, there exist finite positive constants C_1 and C_2 such that, for $0 < t < 1$, and $k = 1, 2$*

$$0 < C_1 t^{\alpha-1} \leq w_k(t) \leq C_2 t^{\alpha-1}$$

where $\alpha = \frac{-\log \gamma}{\log \rho}$ *and* γ *is as in Theorem 3.*

Proof. The lower bound is obvious and rather arbitrary. First let $v^{(0)}$ and $u^{(0)}$ be the left and right eigenvectors of A associated with γ. Then

$$\kappa_0 \left(P_1(1,0) + P_2(1,0)\right) + \kappa_1 \left(P_1(0,1) + P_2(0,1)\right) = \gamma$$

where $\kappa_0 = v_1^{(0)}/v_1^{(0)} + v_2^{(0)}$ and $\kappa_1 = v_2^{(0)}/v_1^{(0)} + v_2^{(0)}$. Using the fact that $f(\Psi(x)) = \Psi(\rho x)$, it is easy to see that

$$w_k(t) = P \sum_{i,j} P_k(i,j) w_{i,j}(t\rho) \quad \text{for } k = 1,2$$

where $w_{i,j}(t)$ is the convolution of densities of $(i+j)$ independent random variables, with i of them identically distributed as $W^{(1)}$ and j of them identically distributed as $W^{(2)}$. Thus

$$\begin{aligned} w_1(t) &\geq \rho\{P_1(1,0)w_1(t\rho) + P_1(0,1)w_2(t\rho)\} \\ &\geq \rho\{(P_1(1,0) \wedge P_1(0,1))(w_1(t\rho) + w_2(t\rho))\} \\ &\geq \rho C_0(\kappa_0 w_1(t\rho) + \kappa_1 w_2(t\rho)) \end{aligned}$$

where $C_0 = P_1(1,0) \wedge P_1(0,1) \wedge (\kappa_0 \wedge \kappa_1)^{-1}$. Let $C_1 = \inf\{\frac{w_1(t)}{t^{\alpha-1}} : \frac{1}{\rho} \leq t \leq 1\}$ and $C_2 = \inf\{\frac{w_2(t)}{t^{\alpha-1}} : \frac{1}{\rho} \leq t \leq 1\}$.

Since $w_i(t)$ is continuous and positive on the positive reals C_1 and $C_2 > 0$. Let n be such that

$$\frac{1}{\rho} \leq \rho^n t < 1.$$

Note that

$$
\begin{aligned}
&\kappa_0 w_1(t) + \kappa_1 w_2(t) \\
&\qquad \ge \rho\,[\kappa_0(P_1(1,0)w_1(\rho t) + P_1(0,1)w_2(\rho t)) \\
&\qquad +\kappa_1(P_2(1,0)w_1(\rho t) + P_2(0,1)w_2(\rho t))] \\
&\qquad = \rho\,[(\kappa_0 P_1(1,0) + \kappa_1 P_2(1,0))w_1(\rho t) + (\kappa_0 P_1(0,1) + \kappa_1 P_2(0,1))w_2(\rho t)] \\
&\qquad = \rho\,[\gamma\kappa_0 w_1(\rho t) + \gamma\kappa_1 w_2(\rho t)] \\
&\qquad = (\rho\gamma)\,[\kappa_0 w_1(\rho t) + \kappa_1 w_2(\rho t)]
\end{aligned}
$$

Iterating the above inequality, we have

$$
\begin{aligned}
\kappa_0 w_1(t) + \kappa_1 w_2(t) &\ge (\rho\gamma)^n(\kappa_0, w_1(\rho^n t) + \kappa_1 w_2(\rho^n t)) \\
&\ge (\rho\gamma)^n(\kappa_0 C_1(\rho^n t)^{\alpha-1} + \kappa_1 C_2(\rho^n t)^{\alpha-1}) \\
&= (\rho\gamma)^n(\kappa_0 C_1 + \kappa_1 C_2)\rho^{n(\alpha-1)}t^{\alpha-1} \\
&= (\kappa_0 C_1 + \kappa_1 C_2)t^{\alpha-1}
\end{aligned}
$$

Using this estimate, we get

$$
\begin{aligned}
w_1(t) &\ge \rho C_0(\kappa_0 C_1 + \kappa_1 C_2)(\rho t)^{\alpha-1} \\
&= \rho^\alpha C_0(\kappa_0 C_1 + \kappa_1 C_2)t^{\alpha-1} \\
&= \left(\frac{C_0}{\gamma}\right)(\kappa_0 C_1 + \kappa_1 C_2)t^{\alpha-1}.
\end{aligned}
$$

A similar estimate prevails for $w_2(t)$ from which the lower bound follows. For the upper bound, we start with the following proposition which may be of independent interest.

Proposition 1. *Let $k \le \alpha < k+1$. Then for every integer $j \in [0,k]$, the jth derivative of w_p exists on $(0,\infty)$ and*

$$
\frac{d^j w_p(t)}{dt^j} = \lim_{N\to\infty} \frac{1}{2\pi} \int_{-N}^{N} (is)^j \Psi_p(-it)dt \qquad \text{for } t > 0 \text{ and } p = 1, 2.
$$

Proof. We shall prove the proposition only for the case $p = 1$. Let

$$\mathbb{K} = \{\Psi_1(-ix): \quad 1 \le x \le \rho\}.$$

Then $\mathbb{K}$ is a compact subset of the unit ball in $\mathbb{C}$, and hence from Theorem 3 there exists a constant A_0 such that

$$|\ f^{(n)}(z)\ | \ \le A_0\gamma^n$$

and

$$|\ D_k(n,z)\cdot 1\ | \ \le A_0\gamma^n$$

where $D_k(n,z)$ is a matrix of partial derivatives of $f_k^{(n)}(z)$. Also using the inequality (for a continuous function h on $[a,b]$ and $|y| \ne 0$)

$$(2) \qquad |\int_a^b h(t)e^{iyt}dt| \le H\cdot\frac{\pi}{|y|} + \frac{b-a}{2}w_h\left(\frac{\pi}{|y|}\right)$$

where

$$H = \sup\{|h(t)|: a \le t \le b\}$$
$$\text{and } w_h(\delta) = \sup\{|h(t_1) - h(t_2)|: a \le t_1 \le t_2 \le b, t_2 - t_1 \le \delta\},$$

we have the estimate

$$|\Psi_k(-ix) - \Psi_k(-iy)| \le A_\beta|y-x|^\beta \text{ for } 1 \le x \le y \le \rho, \beta \in (0,1), \text{ and } k = 1,2.$$

For $t > 0$

$$|\int_1^\rho (iu)^j f_1^{(n)}(\Psi(iu))e^{i\rho^n ut}du| \le \pi(\rho^n t)^{-1}\rho^j A\gamma^n + \frac{(\rho-1)}{2}w_h\left(\frac{\pi}{\rho^n t}\right)$$

where $h(u) = (iu)^j f_1^{(n)}(\Psi(iu))$. Now, $w_h\left(\frac{\pi}{\rho^n t}\right)$ is the same as

$$\sup\{|(iu_1)^j f_1^{(n)}(\Psi_1(iu_1)) - (iu_2)^j f_1^{(n)}(\Psi_1(iu_2))|: 1 \le u_1 \le u_2 \le \rho,\ u_2 - u_1 \le \frac{\pi}{\rho^n t}\}$$
$$\le j\rho^{j-1}A\gamma^n\pi(\rho^n t)^{-1} + \rho^j A\gamma^n A_\beta(\pi\rho^{-n}t^{-1})^\beta$$

where the last inequality follows from the mean value theorem. Combining with the previous estimate, there exists a constant B_β such that

$$\left|\int_1^\rho (ix)^j f_1^{(n)}(\Psi(ix))e^{i\rho^n xt}dx\right| \leq B_\beta(\rho^{-\alpha-\beta})^n max(t^{-1}, t^{-\beta}).$$

Choosing β sufficiently close to 1 so that $\rho^{k+1-\alpha-\beta} < 1$, we have

$$\sum_{n\geq 0}\left|\int_{\rho^n}^{\rho^{n+1}} (ix)^j \Psi_1(-ix)e^{ixt}dx\right| \leq B_\beta \frac{max(t^{-1}, t^{-\beta})}{1-\rho^{k+1-\beta-\alpha}}.$$

Hence

$$w_1(t,j) = \lim_{N\to\infty} \frac{1}{2\pi}\int_{-N}^{N} (ix)^j \Psi_1(-ix)e^{ixt}dx$$

exists and equals

$$\frac{1}{2\pi}\int_{-1}^{1}(ix)^j\Psi_1(-ix)e^{ixt}dx + \sum_{n\geq 0}\frac{1}{2\pi}\int_{\rho^n}^{\rho^{n+1}} \left[(ix)^j\Psi_1(-ix)e^{ixt} + (-ix)^j\Psi_1(ix)e^{-ixt}\right] dx$$

Also for $j \in \{1, \ldots, k\}$ and $0 \notin (a, b)$ it is easy to see that

$$\text{(i)} \int_a^b w(t,j)dt = w(b, j-1) - w(a, j-1)$$

and (ii) if f is any continuous function with compact support in $\mathbb{R} - \{0\}$, then

$$\int_{-\infty}^{\infty} f(t)w_1(t,0)dt = \int_{-\infty}^{\infty} f(t)d\nu(t)$$

where $\nu(t) = P(W \leq t)$. Combining the last two statements completes the proof of Proposition 1. □

Proposition 2. *For* $0 \leq j \leq k$

$$w_p^{(j)}(t) = 0(t^{\alpha-1-j}) \quad \text{as } t \to 0^+$$

Proof. From Proposition 1,

$$w^{(j)}(t) = \lim_{N\to\infty} \frac{i}{\pi} \int_{-N}^{N} (ix)^j \Psi_1(-ix) \sin xt dx.$$

Now,

$$\left| \int_{-1}^{1} (ix)^j \Psi_1(-ix) \sin xt dx \right| \leq t \leq t^{\alpha-1-j} \text{ for } 0 < t \leq 1, \text{ which}$$

implies that $\displaystyle\sum_{n:\ \rho^n < \frac{1}{t}} \left| \int_{\rho^n}^{\rho^{n+1}} (is)^j \Psi_1(-ix) \sin xt dx \right| \leq A\rho^j \rho^2 t^{\alpha-1-j}.$

By using (2) and choosing β sufficiently close to 1, we have

$$\begin{aligned}
&\sum_{n:\ \rho^n \geq \frac{1}{t}} \left| \int_{\rho^n}^{\rho^{n+1}} (ix)^j \Psi_1(-ix) \sin xt dx \right| \\
&\qquad \leq \pi A\rho^j t^{-1} \left(\sum_{n:\ \rho^n \geq \frac{1}{t}} (\rho^j \gamma)^n \right) \\
&\qquad + \sum_{n:\ \rho^n \geq \frac{1}{t}} \left(\frac{\rho-1}{2}\right) \rho^{(j+1)^n} \left[\pi j A \rho^{j-1} \left(\frac{\gamma}{\rho}\right)^n |t|^{-1} + AA_\beta \rho^j \gamma^n \left(\frac{\pi}{t\rho^n}\right)^\beta \right] \\
&\qquad \leq D_\beta t^{\alpha-1-j}
\end{aligned}$$

for some constant D_β, completing the proof of Proposition 2.

Finally the proof of Theorem 3 follows by taking $j = 0$ in Proposition 2. □

Theorem 5 is a refinement of Theorem 4 and is in the spirit of Theorem 2.

Theorem 5. *There exist continuous, positive, multiplicatively periodic functions V_k with period ρ such that as $x \downarrow 0$*

$$x^{1-\alpha} w_k(x) = V_k(x) + 0(1) \quad \textit{for } k = 1, 2.$$

Proof. The proof essentially follows Biggins and Bingham with obvious modifications for the multitype case. We present the proof only to make the paper self contained.

From Theorem 4, it follows that there exists constants C_1 and C_2 such that

$$C_1 \leq \underline{\lim} x^{1-\alpha}(u \cdot w(x)) \leq \overline{\lim} x^{1-\alpha}(u \cdot w(x)) \leq C_2$$

where $w(x) = (w_1(x), w_2(x))$. Let $V_n^{(i)}(x) = \left(\frac{x}{\rho^n}\right)^{1-\alpha} w_i\left(\frac{x}{\rho^n}\right)$ which is bounded as $n \to \infty$. We shall show that $u^{(1)}.\ V_n(x)$ converges where $V_n(x) = (V_n^{(1)}(x), V_n^{(2)}(x))$. Note that

$$\frac{1}{\rho} w_k\left(\frac{x}{\rho^{n+1}}\right) = P_k(1,0) w_1\left(\frac{x}{\rho^n}\right) + P_k(0,1) w_2\left(\frac{x}{\rho^n}\right) + \sum_{i,j} P_k(i,j) w^{i,j}\left(\frac{x}{\rho^n}\right)$$

for $k = 1, 2$. Hence $\frac{\kappa_0}{\rho}\left(\frac{\rho^{n+1}}{x}\right)^{1-\alpha} V_{n+1}^{(1)}(x)$ is the same as

$$\kappa_0 P_1(1,0)\left(\frac{\rho^n}{x}\right)^{1-\alpha}\left[V_n^{(1)}(x) + V_n^{(2)}(x)\right] + \kappa_0 \sum_{i,j} P_1(i,j) w^{(i,j)}(x/\rho^n)$$

and $\frac{\kappa_1}{\rho}\left(\frac{\rho^{n+1}}{x}\right)^{1-\alpha} V_{n+1}^{(2)}(x)$ is the same as

$$\kappa_1 P_2(1,0)\left(\frac{\rho^n}{x}\right)^{1-\alpha}\left[V_n^{(1)}(x) + V_n^{(2)}(x)\right] + \kappa_1 \sum_{i,j} P_2(i,j) w^{(i,j)}(x/\rho^n)$$

Let

$$A_n(x) = \kappa_0 V_n^{(1)}(x) + \kappa_1 V_n^{(2)}(x).$$

Then, after some simple calculations, we have

$$\rho^{-\alpha}\left(\frac{\rho^n}{x}\right)^{1-\alpha} A_{n+1}(x) = \gamma\left(\frac{\rho^n}{x}\right)^{1-\alpha} A_n(x) + \sum\left(\kappa_0 P_1(i,j) w^{i,j} + \kappa_1 P_2(i,j)\right) w^{i,j}(x/\rho^n)$$

and hence

$$\gamma\left(\frac{\rho^n}{x}\right)^{1-\alpha}\left(A_{n+1}(x) - A_n(x)\right) = \sum_{i,j}\left(\kappa_0 P_1(i,j) + \kappa_1 P_2(i,j)\right) w^{i,j}\left(x/\rho^n\right)$$

Thus A_n is a bounded increasing sequence and hence it converges. Let the limit be $A(x)$. Clearly $A(\cdot)$ is multiplicatively periodic. We shall show that $A(\cdot)$ is continuous and that the convergence is uniform in x.

Let $C > C_2$ and $x_0 < \xi$ be so small that $w_k(x) \le Cx^{\alpha-1} < \delta^* < 1$ for $x \le x_0$. Then by induction

$$w^{i,j}(x) \le C^{i+j} x^{(i+j)\alpha-1} \text{ for } x \le x_0$$

Hence,

$$\begin{aligned} A_{n+1}(x) - A_n(x) &\le \frac{1}{\gamma}\sum_{i,j}(\kappa_0, P_1(i,j) + \kappa_1 P_2(i,j))C^{i+j}\left(\frac{x}{\rho^n}\right)^{1-\alpha}\left(\frac{x}{\rho^n}\right)^{(i+j)\alpha-1} \\ &\le \frac{C}{\rho^{n\alpha}}\sum_{i,j}(\kappa_0 P_1(i,j) + \kappa_1 P_2(i,j))(Cx^{\alpha})^{(i+j)-1} \end{aligned}$$

which converges to 0 geometrically. Hence A_n converges uniformly to A for $x \le x_0$. Finally

$$\sup_{x \le x_0\rho^{-n}} \left|A(x) - x^{1-\alpha}(\kappa_0 w_1(x) + \kappa_1 w_2(x))\right| = \sup_{x \le x_0} |A(x) - A_n(x)| \to 0$$

completing the proof of the theorem. $\square$

4. References.

1. Athreya, K. B. and Vidyashankar, A. N. (1994): *Large deviation rates in Branching Processes - II, The multi-type case.*, submitted.

2. Biggins, J. D. and Bingham, N. H. (1990): *Near Constanstancy phenomena in branching processes*, Math. Proc. Camb. Phil. Soc. **110**, 545-558.
3. Biggins, J. D. and Bingham, N. H. (1993): *Large deviations in the supercritical branching process, advances in applied probability* **25 No.** 4, 757-772.
4. Bingham, N. H. (1980): *On the limit of supercritical branching process*, J. Appl. Probab. **25A**, 215-228.
5. Dubuc, S. (1971): *La densité de loi-limite d'un processus en Cascade expansif.*, Z.wahischeinlichkeitsth. Verw. Geb., 281-290.
6. Ellis, Richard S. (1985): *Entropy, Large Deviations, and Statistical Mechanics*, Springer-Verlag, New York.
7. Harris, T. E. (1948): *Branching Processes*, Ann. Math. Statist. **41**, 474, 494.

LARGE DEVIATION RATES FOR BRANCHING PROCESSES - III, THE AGE-DEPENDENT CASE

A PAPER SUBMITTED TO ANNALS OF APPLIED PROBABILITY
ANAND N. VIDYASHANKAR

Departments of Mathematics and Statistics
Iowa State University
Ames, IA 50011

Abstract. In this paper we study the large deviation problem associated with the convergence of empirical age distribution to the stable age distribution. Three situations based on the tails of G arise. We show that under (i) the Malthusian Case the rate is exponential (ii) the sub-exponential case the rate is $1 - G(\cdot)$ and (iii) the super-exponential case the rate is super exponential. Some open problems are indicated.

1. Introduction. Let $\{Z_t : t \geq 0\}$ be an age-dependent branching process with (i) offspring distribution $\{p_j : j \geq 0\}$ and (ii) life-time distribution $G(\cdot)$. Assume that $G(0+) = 0, p_0 = 0, p_1 \neq 1$. Under these conditions, the offspring mean $m \equiv \sum_{j \geq 1} j p_j$, is strictly greater than 1 and the process is supercritical (i.e. the probability of extinction is 0). For details see [3].

For $t \geq 0$ and $x \geq 0$, let $Z_t(x) =$ the number of particles of age less than or equal to x at time t. Athreya and Kaplan [2], Nerman [12], and Kuczek [11] have shown that

$$A(x,t) \equiv \frac{Z_t(x)}{Z_t(\infty)} \to A(x) \text{ with probability 1}$$

as $t \to \infty$ under various moment conditions. This paper is concerned with the large deviation aspects of this convergence. The key idea behind this work is to reduce the large deviation problem (using certain moment conditions) to a study of the decay rates of generating functionals. This leads to three distinct cases based on the tails of $G(\cdot)$, viz. (i) the Malthusian case, (ii) the sub-exponential case, and (iii) the super-exponential case (see Sections [3] and [4] for details). We shall show that under (i) the Malthusian case the rate is exponential (ii) sub-exponential case the rate is $1 - G(\cdot)$, and (iii) the super-exponential case the rate is superexponential. Results related to this work appear in (1), (4), and (5).

The paper is organized as follows: section 2 contains definitions, notations, assumptions and some preliminary results, while section 3 is devoted to the study of decay rates of generating functionals. In sections 4 and 5 we study the large deviation problems and in section 6 we indicate some open problems.

2. Definitions, Notations & Assumptions. Let $(\Omega, \mathcal{F}, P)$ be the underlying probability space and Z_t denote the point process of ages on $[0, \infty)$. It has been noted by Harris ([10], chapter VI-27) that this process is Markovian.

Notations.

N1 $|Z_t(y)|$ is the number of particles at time t, if the process starts at time 0 with one particle of age y. If $y = 0$, we shall denote by $|Z_t|$ the quantity $|Z_t(0)|$.

N2 $Z_t = (x_1, x_2, \dots, x_{|Z_t|})$ is the point process describing the ages of particles alive at time t.

N3 $m = \sum_{j \geq 1} jP_j$.

N4 $Z_t(y; x)$ is the number of particles of ages at most x, if the process starts at time 0 with one particle of age y; we shall denote by $Z_t(x)$ the quantity $Z_t(0; x)$.

N5 $\mathcal{A}$ is the collection of all Borel-measurable functions $s\colon [0, \infty) \to [0, 1]$.

Assumptions.

A1 $p_0 = 0, \quad p_1 \neq 1, \quad 1 < m < \infty.$

A2 $Z_t(y;\cdot)$ has right continuous sample paths.

A3 G is non-lattice and $G(0+) = 0$.

Let $s\colon [0,\infty) \to [0,1]$ be a Borel-measurable function. Define

$$\int s(x) Z_t(y;dx) \triangleq \sum_{j=1}^{|Z_t(y)|} s(x_j)$$

It is shown in [10] that this random integral is a measurable function on Ω. The probability generating functional $\Phi_t[y](s)$ is defined as (as in [9])

$$\Phi_t[y](s) \triangleq E\left[e^{\int \log s(x) Z_t(y;dx)}\right] \tag{1}$$

We recall the following propositions from [10] and [14].

Proposition 2.1. *Φ_t is continuous in the sense that if $s_n \in \mathcal{A}$, $s \in \mathcal{A}$ and $s_n(x) \to s(x)$ for all $x \geq 0$, then $\Phi_t(s_n) \to \Phi_t(s)$ as $n \to \infty$.*

Proposition 2.2. *The map $t \to \Phi_t[y](s)$ is a Borel-measurable function for all $s \in \mathcal{A}$ and all $0 \leq y < \infty$; the map $y \to \Phi_t[y](s)$ is a Borel-measurable function for all $t \geq 0$ and $s \in \mathcal{A}$.*

As in [10], it is easy to show that the functionals Φ_t form a semi-group in t, i.e. for all $0 \leq t, r < \infty$,

$$\Phi_{t+r}(s) = \Phi_t(\Phi_r(s))$$

For $0 \leq y < \infty, \quad 0 \leq t < \infty, \quad 0 < s < 1$, let

$$F_t[y](s) \equiv E(s^{|Z_t(y)|}) \tag{2}$$

denote the probability generating function of $|Z_t(y)|$. One can see that for $s(x) \equiv s$,

$$\Phi_t[y](s) = F_t[y](s).$$

The generating function (2), for $y = 0$, is well-studied in [3]. In fact, the main theorem of the next section is a generalizaton of the rate of convergence of $F_t[0](s)$ to 0 to that of $\Phi_t[0](s)$.

3. Decay Rates for Generating Functionals. Let $p_1 > 0$. Assume that there exists $\gamma > 0$ such that

$$\int_0^\infty e^{\gamma t} dG(t) = \frac{1}{p_1}. \tag{3}$$

γ is called the Malthusian parameter associated with p_1.

Theorem 1 (Malthusian Case). *Assume that the Malthusian parameter γ as defined in (3) above exists. Then under A1 - A3, for $s \in \mathcal{A}$, $\lim_{t\to\infty} e^{\gamma t}\Phi_t[0](s)$ exists, equals $Q(s)$, and satisfies the following:*

$$\begin{aligned} &(i) \quad \text{for every } p > 0 \quad e^{-\gamma p} Q(s) = Q(\Phi_p[\cdot](s)) \\ &(ii) \quad \text{if } s \not\equiv 0, \quad s \not\equiv 1, \quad 0 < Q(s) < \infty. \end{aligned} \tag{4}$$

Before we begin the proof of Theorem 1, we state a proposition whose proof is a part of the standard textbook literature (see [8]).

Proposition 3.1. *Let $H\colon [0,\infty) \to R$ satisfy*

$$H(t) = \xi(t) + \int_0^t H(t-y)dG(y) \tag{5}$$

where ξ is a given bounded Borel-measurable function on $[0,\infty)$ which is directly Riemann integrable (see [3] for definition and sufficient conditions). Then $\lim_{t\to\infty} H(t)$ exists and equals $\frac{1}{\mu}\int_0^\infty \xi(x)dx$ where $\mu \triangleq \int_0^\infty t dG(t)$.

We now start the proof of Theorem 1.

Proof of Theorem 1. The probability generating functional Φ_t satisfies (see also [14])

$$\Phi_t[0](s) = (1 - G(t))s(t) + \int_0^t f(\Phi_{t-u}[0](s))G(du). \tag{6}$$

(6) is an immediate consequence of the branching property if we decompose Ω into the events that either the particle is still alive at time t or it has died at a time $u \leq t$ and if we express $Z_t(0;\cdot)$ in terms of the process generated by the offspring of the initial particle.

One can rewrite (6) as

$$\Phi_t[0](s) = \xi(t) + p_{k_0} \int_0^t \Phi_{t-u}[0](s)G(du) \tag{7}$$

where

$$\xi(t) = \xi_1(t) + \xi_2(t), \text{ with } \xi_1(t) = (1 - G(t))s(t)$$

and

$$\xi_2(t) = \int_0^t (f(\Phi_{t-u}[0](s)) - p_1(\Phi_{t-u}[0](s)))dG(u)$$

Multiplying equation (7) by $e^{\gamma t}$ and setting $\xi^{(\gamma)}(t) = e^{\gamma t}\xi(t)$, $H^{(\gamma)}(t,s) = e^{\gamma t}\Phi_t[0](s)$, and $G^{(\gamma)}(u) = pe^{\gamma u}G(u)$, one obtains

$$H^{(\gamma)}(t,s) = \xi^{(\gamma)}(t) + \int_0^t H^{(\gamma)}(t-u,s)dG^{(\gamma)}(u) \tag{8}$$

Thus to complete the proof of Theorem 1, it is enough (by Proposition 3.1) to show that $\xi^{(\gamma)}(t)$ is directly Riemann integrable. The direct Riemann integrability of $e^{\gamma t}\xi_1(t)$ follows from the existence of the Malthusian parameter. However, we include the proof to make the paper self-contained. Lemma 3.3 establishes the direct Riemann integrability of $e^{\gamma t}\xi_2(t)$.

Lemma 3.1. *$e^{\gamma t}\xi_1(t)$ is directly Riemann integrable.*

Proof. Since $e^{\gamma t}\xi_1(t) \leq e^{\gamma t}(1 - G(t))$, it is enough (by condition (iii) on page 46, of [3]) to establish that $e^{\gamma t}(1 - G(t))$ is directly Riemann integrable. To do this we will show that, for $h > 0$,

$$\text{(i)} \quad \sum_{n \geq o} U(n,h) < \infty$$

and

$$\text{(ii)} \quad h \sum_{n \geq 0} (U(n,h) - L(n,h)) \to 0 \text{ as } h \to 0$$

where $U(n,h) = \sup\limits_{nh \leq t \leq (n+1)h} e^{\gamma t}(1 - G(t))$ and $L(n,h)$ is the corresponding infimum. Under the assumption that the Malthusian parameter exists, $e^{\gamma t}(1 - G(t))$ is integrable; hence

$$\begin{aligned} \infty > \int_0^\infty e^{\gamma t}(1 - G(t))dt &= \sum_{n \geq 0} \int_{nh}^{(n+1)h} e^{\gamma t}(1 - G(t))dt \\ &\geq \sum_{n \geq 0} e^{\gamma nh}(1 - G(t))h \\ &= he^{-h} \sum_{l \geq 1} e^{\gamma lh}(1 - G(lh)) \\ &= he^{-h}A(h) \end{aligned} \tag{9}$$

where

$$A(h) = \sum_{l \geq 1} e^{\gamma lh}(1 - G(lh))$$

Similarly one can show that

$$\int_0^\infty e^{\gamma t}(1 - G(t))dt \leq e^h(A(h) + 1) \tag{10}$$

Combining (9) and (10) one has

$$he^{-h}A(h) \le \int_0^\infty e^{\gamma t}(1 - g(t))dt \le e^h(A(h) + 1) \tag{11}$$

Now, note that

$$e^{-h}A(h) \le \sum_{n\ge 0} U(n, h) \le e^h(A(h) + 1)$$

and

$$e^{-h}A(h) \le \sum_{n\ge 0} L(n, h) \le e^h(A(h) + 1).$$

The last two inequalities along with (9) establish the lemma.

Lemma 3.2 is technical and is needed for establishing the direct Riemann integrability of $e^{\gamma t}\xi_2(t)$.

Lemma 3.2. *For any* $\gamma' < \gamma$,

$$\sup_{t\ge 0} e^{\gamma' t}|\Phi_t[0](s)| = K_{\gamma'} < \infty \tag{12}$$

for $s \in \mathcal{A}, s \not\equiv 1$.

Proof. Given $\gamma' < \gamma$, let $1 > p' > p_1$ be such that

$$\int_0^\infty e^{\gamma' t} dG(t) = \frac{1}{p'}.$$

(Note that such a p' exists by continuity of the Laplace transform on the domain of its finiteness). Further, by the mean value theorem (since $\Phi_t[0](s) \to 0$ as $t \to \infty$) it follows that there exists a $t_0(s)$ such that for $t \ge t_0$,

$$f(\Phi_t[0](s)) \le p'\Phi_t[0](s) \tag{13}$$

Hence

$$\Phi_t[0](s) \le (1 - G(t - t_0)) + p' \int_0^t \Phi_{t-u}[0](s)dG(u)$$

The lemma follows (using Lemma 3.1) by a comparison argument from Proposition 3.1.

Lemma 3.3. *$e^{\gamma t}\xi_2(t)$ is directly Riemann integrable.*

Proof. The estimate

$$|f(\Phi_t[0](s)) - p_{k_0}\Phi_t[0](s)| \leq C\Phi_t[0](s) \leq Ce^{-\gamma' t}$$

follows from

$$|f(x) - p_1 x| \leq Cx \text{ for } x \in [0,1]$$

(where C is a finite positive constant) and Lemma 3.2 where γ' is such that $\gamma < \gamma' < 2\gamma$. Thus,

$$\xi_2(t) = \int_0^t [f(\Phi_{t-u}[0](s) - p_{k_0}\Phi_{t-u}[0](s)]\, dG(u)$$

$$\leq C \int_0^t e^{-2\gamma'(t-u)} dG(u) \tag{14}$$

Hence $e^{\gamma t}\xi_2(t)$ is bounded above by $B(t) = e^{\gamma t}\int_0^t e^{-2\gamma'(t-u)} dG(u)$. Whence, to complete the proof of the lemma, it is enough to show that $B(t)$ is directly Riemann integrable (using (iii) on page 146 of [3]). To do this, we will show that (using (i) on page 146 of [3])

(i) $B(\cdot)$ is bounded and

(ii) $\sum_{n\geq 1} \sup_{n\leq t\leq (n+1)} B(t) < \infty$.

One can show that if $\int B(t)dt < \infty$ then (i) and (ii) follow. The integrability of $B(t)$ can be seen using Tonelli's theorem (see [13]) as follows:

$$
\begin{aligned}
\int_0^\infty B(t)dt &= \int_0^\infty e^{\gamma t} \int_0^t e^{-2\gamma'(t-u)} dG(u)dt \\
&= \int_0^\infty \left[\int_0^t e^{(\gamma-2\gamma')(t-u)} e^{\gamma u} dG(u) \right] dt \\
&= \int_0^\infty \left[\int_u^\infty e^{(\gamma-2\gamma')(t-u)} dt \right] e^{\gamma u} dG(u) \\
&= \left(\int_0^\infty e^{\gamma u} dG(u) \right) \left(\int_0^\infty e^{(\gamma-2\gamma')t} dt \right) \\
&< \infty \quad \text{(by the choice of } \gamma').
\end{aligned}
$$

□

An immediate consequence of Theorem 1 is the following:

Corollary 1. $\overline{\lim_{t\to\infty}}\, e^{\gamma t}\Phi_t[y](s) \leq C(s,y)$ *where C is a finite positive constant depending on s and y.*

Remark 1. Theorem 1 (above) is a generalization of Theorem IV.7.1 of [3] in two directions. First, it is extended to generating functionals thus extending the domain of applicability. Second, it also considers the situation when $p_1 = 0$ (see also Remark 3).

Theorem 2 below considers the decay rates for generating functionals when G has "sub-exponential" tails. First, we recall the following definition from [3].

Definition. The "sub-exponential" class $\mathcal{L}$ consists of all distribution functions G such that

$$\lim_{t\to\infty} \frac{1 - G^{2*}(t)}{1 - G(t)} = 2 \tag{15}$$

Theorem 2 (sub-exponential case). *Let $G \in \mathcal{L}$ and $k_0 = \inf\{k \geq 1 \colon p_k > 0\}$. Let $l_s = \inf_{t \geq 0} s(t) \leq \sup_{t \geq 0} s(t) = U_s$. Then under A1 - A3,*

$$\frac{ls}{1-p} \leq \varliminf_{t \to \infty} \frac{\Phi_t[0](s)}{1-G(t)} \leq \varlimsup_{t \to \infty} \frac{\Phi_t[0](s)}{1-G(t)} \leq \frac{U_s}{1-p}$$

where $p = p_{k_0}$ and $s \in \mathcal{A}, s \not\equiv 1$.

Proof. Note that from (6), we have

$$\Phi_t[0](s) = s(t)(1-G(t)) + \int_0^t f(\Phi_{t-u}[0](s))dG(u)$$

since $\Phi_t[0](s) \to 0$ as $t \to \infty$, there exists a $t_0(s)$ such that $\forall t \geq t_0$

$$(p-\varepsilon)\Phi_t[0](s) \leq f(\Phi_t[0](s)) \leq (p+\varepsilon)\Phi_t[0](s).$$

Hence

$$\Phi_t[0](s) \leq s(t)(1-G(t)) + r(t) + (p+\varepsilon)\int_0^t \Phi_{t-u}[0](s)dG(u)$$

where

$$r(t) = \int_{t-t_0}^t f(\Phi_{t-u}[0](s))dG(u) - (p+\varepsilon)\int_{t-t_o}^t \Phi_{t-u}[0](s)dG(u).$$

Since

$$C_2(G(t) - G(t-t_0)) \leq r(t) \leq C_1(G(t) - G(t-t_0)),$$

it follows that

$$R_2(t) + (p-\varepsilon)(x * G)(t) \leq x(t) \leq R_1(t) + (p+\varepsilon)(x * G) \tag{16}$$

where $x(t) = \Phi_t[0](s)$, $R_1(t) = s(t)(1-G(t)) + C_1(G(t) - G(t-t_0))$, and $R_2(t) = s(t)(1-G(t)) + C_2(G(t) - G(t-t_0))$.

Iterating (16), one gets

$$R_2 * V_{(p-\varepsilon)}(t) \le x(t) \le R_1 * V_{p+\varepsilon}(t) \tag{17}$$

where

$$V_{p\pm\varepsilon}(t) = \sum_{n\ge 0}(p \pm \varepsilon)^n G^{n*}(t)$$

and G^{n*} is the n-fold convolution of G with itself.

To complete the proof of the theorem, (since ε is arbitrary) we establish the following:

(i) $\displaystyle \lim_{t\to\infty} \frac{((G - G(\cdot - t_0)) * V_p)(t)}{1 - G(t)} = 0$

(ii) $\displaystyle \frac{l_s}{1-p} \le \overline{\lim_{t\to\infty}} \frac{(s(1-G) * V_p)(t)}{1-G(t)} \le \underline{\lim} \frac{(s(1-G) * V_p)(t)}{1-G(t)} \le \frac{U_s}{1-p}$

The result would then follow from the definition of R_1 and R_2. We start with the proof of (i). The proof is based on the tail behavior of the renewal function

$$V_\gamma(t) = \sum_{n\ge 0} \gamma^n G^{n*}(t).$$

The next proposition describes the tail behavior of $V_\gamma(\cdot)$ and the proof is available in [3].

Proposition 3.2. *Let $G \in \mathcal{L}$ and $0 < \gamma < 1$, then*

$$\lim_{t\to\infty} \frac{(1-\gamma)^{-1} - V_\gamma(t)}{1 - G(t)} = \frac{\gamma}{1-\gamma^2} \tag{18}$$

Proof. See [3], page 150 (Theorem IV.4.3).

Now to complete the proof of (i), we rewrite it as

$$\frac{(G * V_p)(t)}{1-G(t)} - \frac{(G * V_p)(t-t_0)}{1-G(t-t_0)}\frac{(1-G(t-t_0))}{1-G(t)} + \frac{(G(\cdot - t_0) * (V_p - V_p(\cdot - t_0)))(t)}{1-G(t)}.$$

Since $G \in \mathcal{L}$, by a lemma of Chistyakov (see [3], page 148)

$$\lim_{t\to\infty} \frac{1-G(t-t_0)}{1-G(t)} = 1 \quad \forall \quad 0 \le t_0 < \infty;$$

and from Proposition 3.2 one can show that

$$\lim_{t\to\infty} \frac{V_p(t)-V_p(t-t_0)}{1-G(t)} = 0. \text{ Thus (i) follows.}$$

Finally, we consider (ii). Now from Theorem IV.5.3B of [3] it follows that

$$\frac{l_s}{1-p} \le \overline{\lim_{t\to\infty}} \frac{(s(t)(1-G(t)) * V_p(t)}{1-G(t)} \le \underline{\lim} \frac{(s(t)(1-G(t)) * V_p(t)}{1-G(t)} \le \frac{U_s}{1-p}$$

completing the proof of Theorem 2. □

Our next theorem considers the case when G has superexponentially decaying tails and is concerned with the generating function of $|Z_t|$. (Note that s is not a function in the next theorem.)

Theorem 3 (super-exponential case). *Assume that*

$$1 - G(x) \le C_1 e^{-\theta e^{-\alpha x}}, \quad x > 0$$

where C_1, and θ are finite positive constants. Assume further that $p_0 = 0 = p_1$ and $p_2 > 0$. Let α be the Malthusian parameter associated with 2 (the minimum family size), i.e.

$$\int_0^\infty e^{\alpha x} dG(x) = \frac{1}{2}$$

Let the number of births $N(t)$ in $(0,t)$ satisfies the following condition:

$$\log P(N(t) \le Kt) = 0(e^{\alpha t})$$

where K is a large positive constant. There exists a constant $K_1(\theta)$ such that

$$\log s \le \overline{\lim}\, e^{-\alpha t} \log F_t[0](s) \le \underline{\lim}_{t\to\infty}\, e^{-\alpha t} \log F_t[0](s) \le K_1 \log(s+1) \quad \textit{for } 0 \le s \le 1.$$

Proof. The proof of the lower bound is easily seen as follows:

$$F_t[0](s) = E(s^{|Z_t|}) \geq E\left(s^{|Z_t|} \colon |Z_t| \geq e^{\alpha t}\right) \geq s^{e^{\alpha t}} P(|Z_t| \geq e^{\alpha t})$$

Thus

$\varliminf_{t\to\infty} e^{-\alpha t} \log F_t[0](s) \geq \log s + \varlimsup_{t\to\infty} e^{-\alpha t} \log P(|Z_t| \geq e^{\alpha t}) = \log s$ (by the choice of α).

Let $N(t)$ denote the number of births in $(0,t)$. Then

$$\begin{aligned} E(s^{|Z_t|}) &= E(s^{|Z_t|} \colon N(t) = 0) + E(s^{|Z_t|} \colon N(t) \geq 1) \\ &= s(1 - G(t)) + E(s^{|Z_t|} \colon N(t) \geq 1) \\ &\leq s(1 - G(t)) + E(s^{2^{N(t)-1}} \colon N(t) \geq 1) \\ &\leq s e^{-\theta e^{\alpha t}} + \sum_{k\geq 1} s^{2^{(k-1)}} P(N(t) = k) \\ &= s e^{-\theta e^{\alpha t}} + \left(\sum_{k=1}^{j(t)} + \sum_{k \geq j(t)+1} \right) s^{2^{(k-1)}} P(N(t) = k) \\ &\leq s e^{-\theta e^{\alpha t}} + s^{2^{j(t)}} + P(N(t) \leq j(t)) \end{aligned}$$

where $j(t) = \left[\frac{ln\theta}{ln_2} + \frac{\alpha}{ln_2} t\right]$. For such a choice of $j(t)$ we have

$$\begin{aligned} F_t[0](s) &\leq e^{-\theta e^{\alpha t}} \left(s + 1 + P(N(t) \leq j(t)) e^{\theta e^{\alpha t}} \right) \\ \varlimsup_{t\to\infty} e^{-\alpha t} \log F_t[0](s) &\leq -\theta + \log(s+1) \text{ (by assumption)} \end{aligned}$$

thus completing the proof of our theorem. □

4. Large deviations. In this section we shall derive the large deviation rates for the convergence of empirical age distribution (viz., $(|Z_t|)^{-1} Z_t(x)$) to the stable age distribution $A(x)$ where $A(x) = \left[\int_0^\infty e^{\alpha u}(1 - G(u)) du\right]^{-1} \left[\int_0^x e^{\alpha u}(1 - G(u)) du\right]$ and α is the root of the equation

$$m \int_0^\infty e^{\alpha u} dG(u) = 1 \tag{19}$$

α is called the Malthusian parameter associated with m.

Theorem 4. *Let $k_0 = \inf\{k \geq 1\colon p_k = 0\}$. Let γ be the Malthusian parameter associated with p_{k_0} (see equation (3) above). Assume that $E(e^{\theta|Z_{r_0}|}) < \infty$ for some θ and a suitable choice of r_0 (to be specified in the proof of the theorem). Then for every $\varepsilon > 0$ and $x \geq 0$*

$$\lim_{t\to\infty} e^{-\gamma t} P\left(|\frac{Z_t(x)}{|Z_t|} - A(x)| > \varepsilon\right)$$

exists and is finite and positive.

Before proving Theorem 4 we shall make two observations that are crucial to the proof of Theorem 4. We begin by setting $Y_r(y) = Z_r(y;x)-(A(x)+\varepsilon)|Z_r(y)|$, $A_r(\theta,y) = E(e^{\theta Y_r(y)})$ and $B_r(y) = E(-Y_r(y))$.

(i) *There exists $r_1 > 0$ such that*

$$\inf_{y\geq 0} B_{r_1}(y) > 0$$

(ii) $\sup_y |\frac{log A_{r_1}(\phi,y)}{\phi} - B(y)| \to 0$ as $\phi \to 0$.

(i) is an obvious consequence of Lemma 1 of [2]. To prove (ii) we begin with the following observations:

(a) *$Y_r(y)$ is a uniformly integrable collection of random variables.*

(b) *$Y_r(y)e^{\phi Y_r(y)}$ is a uniformly integrable collection of random variables.*

(a) and (b) can be seen, for example, by choosing $r_0 > r_1$ and $\phi = \frac{\theta}{2}$ in the theorem.

Finally, to prove (ii) it is enough to show that

$$\sup_{y\geq 0} |\frac{A'_{r_1}(\phi,y)}{A_{r_1}(\phi,y)} - B(y)| \to 0 \quad \text{as } \phi \to 0.$$

Let

$$\underline{A}_{r_1}(\phi) = \inf_{y\geq 0} Ar_1(\phi,y)$$

and

$$\overline{A}_{r_1}(\phi) = \sup_{y \geq 0} A_{r_1}(\phi, y)$$

Then it is easy to see that

$$0 < \lim_{\phi \to 0} \underline{A}_{r_1}(\phi) < \overline{\lim_{\phi \to 0}}\, \overline{A}_{r_1}(\phi) < \infty$$

Hence

$$\sup_{y \geq 0} \left| \frac{A'_{r_1}(\phi, y) - B(y) A_{r_1}(\phi, y)}{A_{r_1}(\phi, y)} \right| \leq (\underline{A}_{r_1}(\phi))^{-1} \sup_{y \geq 0} |A'_{r_1}(\phi, y) - B(y) A_{r_1}(\phi, y)|$$

Finally (ii) follows from (a) and (b). We now proceed to prove the theorem.

Proof of Theorem 4. Let r_0 be such that the conditions (i) and (ii) above are true. For $t \geq 0$, let $\mathcal{F}_t = \sigma\langle Z_s \colon 0 \leq s \leq t\rangle$ be the σ-algebra generated by the process up to time t. Set $A(x,t) = (|Z_t|)^{-1} Z_t(x)$. Then, conditioned on $\mathcal{F}_{t-r_0}, A(x,t) > A(x) + \varepsilon$ if and only if

$$\sum_{i=1}^{|Z_{t-r_0}|} Y_{r_0}(x_i) > 2 \sum_{i=1}^{|Z_{t-r_0}|} B_r(x_i)$$

where $Y_{r_0}(x_i)$ are independent random variables. Thus

$$P(A(x,t) > A(x) + \varepsilon | \mathcal{F}_{t-r_0}) \leq \prod_{i=1}^{|Z_{t-r_0}|} A_{r_0}(\theta, x_i) e^{-2\theta B_{r_0}(x_i)}$$

$$= e^{\left(\sum_{i=1}^{|Z_{t-r_0}|} log \left[(A_{r_0}(\theta, x_i))^{\frac{1}{\theta}} e^{-2B_{r_0}(x_i)} \right]^{\theta} \right)}$$

By (i) and (ii) there exists a $\theta_0 > 0$ such that $0 < \left[(A_{r_0}(\theta_0, y))^{\frac{1}{\theta_0}} e^{-2B_{r_0}(y)} \right]^{\theta_0} < 1$ for all $y \geq 0$. Hence,

$$(20) \qquad e^{\gamma t} P(A(x,t) > A(x) + \varepsilon) \leq E\left[e^{\int logh(y) dZ_{t-r_0}(y)} \right] = e^{\gamma t} \Phi_{t-r_0}[0](h)$$

where

$$h(y) = \left[(A_{r_0}(\theta_0, y))^{\dfrac{1}{\theta_0}} e^{-2B_{r_0}(y)} \right]^{\theta_0}$$

Thus

$$e^{\gamma t} P\left(\sum_{i=1}^{|\tilde{x}|} Y_{r_0}(x_i) > 2\sum_{i=1}^{|\tilde{x}|} B_{r_0}(x_i)\right) P((Z_{t-r_0} = \tilde{x}))$$

$$\leq e^{\gamma t} \sum_{e^{i=1}}^{|\tilde{x}|} log h(x_i) P(Z_{t-r_0} = \tilde{x})$$

Now, letting $t \to \infty$, we see that the LHS converges to

$$P\left(\sum_{i=1}^{|\tilde{x}|} Y_{r_0}(x_i; x) > 2\sum_{i=1}^{|\tilde{x}|} B_{r_0}(x_i)\right) Q(s_0)$$

where s_0 is the function defined by

$$\begin{aligned} s_0(y) &= 1 \quad \text{if } y = x_i \\ &= 0 \quad \text{if } y \neq x_i \end{aligned}$$

and $\tilde{x} = (x_i, x_2, \ldots, x_{|x|})$ and the RHS converges to $\left(\prod_{i=1}^{|\tilde{x}|} h(x_i)\right) Q(s_0)$.

Similar calculations prevail for the other side. Also, by Theorem 1, the RHS of (18) converges to $Q(h)e^{-\gamma r_0}$. Thus by a generalized version of dominated convergence theorem, (see [13], pages 92, 270) applied to the space of types, Theorem 4 follows. □

Remark 4. Unlike the single type and multitype cases (see [1] and [5]), the rate of decay of $P(|A(x,t) - A(x)| > \varepsilon)$, even if $p_1 = 0$, is only exponential . This can be seen intuitively as follows: to obtain superexponential decay it is just not enough to have $p_1 = 0$ but also the life lengths must be small. See also Theorems 3 and 6.

Remark 5. Using Theorem 4 one can show that there does not exist a large deviation principle (see [7]) for the convergence of empirical age distribution to the stable age distribution. The details of the argument are similar to the one in [4].

Theorem 5 (sub-exponential case). *Let $k_0 = \inf\{k \geq 1 \colon p_k > 0\}$ and $G \in \mathcal{L}$. Assume that $E(e^{\theta|Z_{r_0}|}) < \infty$ for some positive θ and a suitable choice of r_0. Then for every $\varepsilon > 0$, there exists positive finite constants $k_1(\varepsilon)$ and $k_2(\varepsilon)$ such that*

$$\frac{k_1(\varepsilon)}{1-p_{k_0}} \leq \varliminf_{t\to\infty}(1-G(t))^{-1}P\left(\frac{Z_t(x)}{|Z_t|} - A(x)| > \varepsilon\right)$$
$$\leq \varliminf_{t\to\infty}(1-G(t))^{-1}P\left(|\frac{Z_t(x)}{|Z_t|} - A(x)| > \varepsilon\right) \leq \frac{k_2(\varepsilon)}{1-p_{k_0}}$$

Proof. Similar to the proof of Theorem 4. □

Theorem 6 (super-exponential case). *Let $k_0 = \inf\{k \geq 1 \colon p_{k_0} > 0\}$. Assume that the conditions of Theorem 3 above hold and that $E(e^{\theta|Z_{r_0}|}) < \infty$ for some positive θ and a suitable choice of r_0. Then for every $\varepsilon > 0$, there exists a finite constant k_1 such that*

$$\varlimsup_{t\to\infty} e^{\gamma t}\log P\left(|\frac{Z_t(x)}{|Z_t|} - A(x)| > \varepsilon\right) \leq k_1$$

Proof. As in the proof of Theorem 4, (see (18) above)

$$P\left(|\frac{Z_t(x)}{|Z_t|} - A(x)| > \varepsilon\right) \leq \Phi_{t-r_o}[0](h).$$

It is possible to show that there exists a θ_0 such that

$$0 < \sup_{y\geq 0} h(y) \leq \eta < 1.$$

For such a choice of θ_0,

$$\Phi_{t-r_0}[0](h) \leq F_{t-r_0}[0](\eta).$$

The result follows from Theorem 3. □

We now turn our attention to weakening the exponential moment hypothesis to a polynomial moment hypothesis. As one should expect, the details of the proof are more complicated but the key idea goes back to [1] and [5].

Theorem 7. *Let $k_0 = \inf\{k \geq 1: p_k > 0\}$ and γ be the Malthusian parameter associated with p_{k_0} (see equation (3) above). Assume further that the residual life time $Gy(x)(= (1-G(y))^{-1}(G(x+y)-G(y)))$ satisfies the following condition: There exists an integer r_0 such that $\inf_{y\geq 0} Gy(\frac{r_0}{2}) = \theta > 0$. Let $r_0 > 0$ and $p > 2$ be such that $M_{r_0}^{(p)} = E(|Z_{r_0}|^p) < \infty$ and $M_{r_0}^{(p)} e^{\gamma r_0} > 1$. Then under A1 - A3, $\lim_{t\to\infty} e^{\gamma t} P\left(|\frac{Z_t(x)}{|Z_t|} - A(x)| > \varepsilon\right)$ exists and is positive and finite.*

Before we begin the proof of the theorem we shall prove two propositions which are crucial to the proof of our main theorem.

Proposition 4.1. *For every $0 \leq s \leq 1$, $e^{\gamma t} F_t[0](s)$ is eventually increasing.*

Proof. Note that there exists a $t_0 > 0$ such that for every $t \geq t_0(s)$

$$p_{k_0} F_{t+u}[0](s) - e^{-\gamma t} F_t[0](s) \geq 0$$

for every $y \geq 0$. Thus, for $t \geq t_0$ and $u > 0$

$$\begin{aligned} & e^{\gamma(t+u)} F_{t+u}[0](s) - e^{\gamma t} F_t[0](s) \\ & = e^{\gamma t}\left[s(1-G(t+u)) + \int_0^{t+u} f(F_{t+u-v}[0](s))dG(v) - F_t[0](s)\right] \\ & \geq e^{\gamma t}\{[s - F_t[0](s)](1-G(t+u)) + (p_{k_0}F_{t-u}[0](s) - F_t[0](s))\, G(t+u)\} \\ & \geq 0 \end{aligned}$$

thus completing the proof of the proposition.

Proposition 4.2. *Assume that there exists an integer r_0 such that $\inf_{y\geq 0} G_y\left(\frac{r_0}{2}\right) = \theta > 0$. Then there exists a generating function $\Psi: [0,1] \to [0,1]$ such that $\Psi'(1) > 1$ and $F_{r_0}[y](s) \leq \Psi(s)$.*

Proof. $F_{r_0}[y](s) = E(s^{|Z_{r_0}(y)|})$. Let L_y denote the random variable with distribution G_y. Then

$$\begin{aligned} F_{r_0}[y](s) &= E\left(s^{|Z_{r_0}(y)|} : L_y \leq \frac{r_0}{2}\right) + E\left(s^{|Z_{r_0}(y)|} : L_y \geq \frac{r_0}{2}\right) \\ &\leq E\left(s^{|\frac{Z_{r_0}}{2}|}\right) G_y(\frac{r_0}{2}) + s(1 - G_y(\frac{r_0}{2})) \end{aligned}$$

Thus,

$$F_{r_0}[y](s) \le (F_{\frac{r_0}{2}}[0](s))Gy(\frac{r_0}{2}) + (1 - Gy(\frac{r_0}{2}))s \equiv h(s)$$

Let

$$\Psi(s) = \theta F_{\frac{r_0}{2}}[0](s) + (1-\theta)s.$$

Note that $\Psi'(1) > 1$ and $h(s) \le \Psi(s)$ thus proving the proposition.

Finally, note that

$$\begin{aligned}(M_{r_0}(y;x) - &(A(x)+\varepsilon)M_{r_0}(y)) \\ &\ge \left(\inf_y M_{r_0}(y;x) - (A(x)+\varepsilon)\sup_y M_{r_0}(y)\right) \\ &= \mu_{r_0}(x)\end{aligned}$$

We are now ready to prove the main theorem.

Proof of Theorem 7. Recall that $A(x,t) = (|Z_t|)^{-1}Z_t(x)$. Then

$$P(A(x,t) > A(x) + \varepsilon) = E\left(P(A(x,t) > A(x) + \varepsilon | \mathcal{F}_t)\right)$$

Now by applying central limit theorem arguments

$$\begin{aligned}P(A(x,t) > A(x) + \varepsilon | \mathcal{F}_{t-r_0}) &= P\left(\sum_{i=1}^{|\tilde{x}|} Y_{r_0}(x_i) > |\tilde{x}|\mu_{r_0}(x) | \mathcal{F}_{t-r_0}\right) \\ &\le \frac{E\left(\sum_{i=1}^{|\tilde{x}|} Y_{r_0}(x_i)\right)^{2k}}{|\tilde{x}|^{2k}(\mu_{r_0}(x))^{2k}} \\ &\le \frac{C(x;\varepsilon)}{|\tilde{x}|^k}\end{aligned}$$

where $C(x;\varepsilon)$ is a finite positive constant. Thus

$$\begin{aligned}P(A(x,t) > A(x) + \varepsilon) &\le C(x,\varepsilon)\int \frac{1}{|\tilde{x}|^k} dP^{t-r_0}(\tilde{x}) \\ &= C\sum_{j\ge 1}\int_{\{\tilde{x}:\ |\tilde{x}|=j\}} \frac{1}{|\tilde{x}|^k} dP^{t-r_0}(\tilde{x}) \\ &= C\sum_{j\ge 1}\frac{1}{j^k}P(|Z_t| = j) \qquad (21)\end{aligned}$$

Using the fact that (see [1], [5] for more details)

$$\Gamma(k)E\left(\frac{1}{|Z_t|^k}\right) = \int_0^\infty F_t[0](e^{-x})x^{k-1}dx \text{ and}$$

setting $e^{-x} = s$, the above integral reduces to $\int_0^1 F_t[0](s)\frac{|\log s|^{k-1}}{s}ds$. Hence, $e^{\gamma t}\int_0^t F_t[0](s)\frac{|\log s|^{k-1}}{s}ds$ converges, by monotone convergence theorem (in lieu of Proposition 5.1) to $\int_0^1 Q(s)K(s)ds$ where $K(s) = \frac{|log s|^{k-1}}{s}$.

From (4) and proposition 5.2

$$\begin{aligned} Q(s) &= e^{\gamma r_0}Q(F_{r_0}[\cdot](s)) \\ &\leq e^{\gamma r_0}Q(\Psi(s)) \end{aligned} \tag{22}$$

Set $g(s) = \Psi^{-1}(s)$ and fix $0 < s_0 < 1$. Let $t_0 = g(s_0)$ and for $k \geq 1$ $t_k = g_k(s_0) = g_{k-1}(g(s_0))$. Consider

$$\int_{t_0}^1 Q(s)K(s)ds = \sum_{k\geq 1}\int_{t_{k-1}}^{t_k} Q(s)K(s)ds.$$

Define $I_k = \int_{t_{k-1}}^{t_k} Q(s)K(s)ds$. Then

$$I_{k+1} = \int_{t_k}^{t_{k+1}} Q(s)K(s)ds \leq \int_{t_k}^{t_{k+1}} e^{\gamma r_0}Q(\Psi(s))K(s)ds.$$

By setting $\Psi(s) = u$ we have

$$\begin{aligned} I_{k+1} &\leq \int_{t_{k-1}}^{t_k} e^{\gamma r_0}Q(u)\frac{K(g(u))}{\Psi'(g(u))}du \\ &= \int_{t_{k-1}}^{t_k} Q(u)K(u)\left[\frac{e^{\gamma r_0}K(g(u))K(u)}{\Psi'(g(u))}\right]du. \end{aligned}$$

Since

$$\frac{e^{\gamma r_0}K(g(u))K(u)}{\Psi'(g(u))} \to \frac{e^{\gamma r_0}}{(\Psi'(1))^p} < 1 \text{ as } u \to 1$$

there exists $1 \geq u > u_0$ such that

$$\frac{e^{\gamma r_0}K(g(u))K(u)}{\Psi'(g(u))} < \left(\frac{e^{\gamma r_0}}{(\Psi'(1))^p} + \eta\right) = \beta < 1$$

Now by choosing $s_0 = u_0$, we have

$$I_{k+1} \leq \beta I_k \quad \forall \ k \geq 1.$$

The result follows from generalized version of dominated convergence theorem just as in Theorem 4. □

5. Rates for $P(W - W_t > \varepsilon)$. For $x \geq 0$, let $A(x)$ be the stationary age distribution and

$$V_y(x) = e^{-\alpha x}(1 - G_y(x))^{-1} \int_x^\infty e^{\alpha u} dG_y(u)$$

$V_y(x)$ is called the reproductive value associated with a particle of age y. When $y = 0$, we denote $V_0(x)$ by $V(x)$.

$$W_t = e^{\alpha t} \sum_{i=1}^{|Z_t|} V(x_i) \tag{23}$$

where $\tilde{x} = (x_1, x_2, \ldots, x_{|Z_t|})$ is the age chart at time t. Then it is known that (see [3]) $\{(W_t, \mathcal{F}_t)\colon t \geq 0\}$ is a non-negative martingale and hence converges with probability 1. Our aim is to investigate the decay rate of $P(W - W_t > \varepsilon)$. In fact, we shall show that this rate is super exponential. We start with the following theorem.

Theorem 8. *Assume that there exists a* $\theta_0 > 0$ *and* $r_0 > 0$ *such that* $E(e^{\theta_0 |Z_{r_0}|}) < \infty$, *and* $\inf_{y \geq 0} V(y) > 0$. *Then there exists a* $\theta^* > 0$ *such that*

$$\sup_{t \geq 0} E(e^{\theta^* W_t}) < \infty$$

Proof. Set $\phi_t(\theta) = E(e^{\theta W_t})$. Let $\{\theta_t : t \geq 0\}$ be such that

$$E(e^{\theta_t W_t}) = E(e^{\theta_s W_s}) = k < \infty$$

for $t > s$. But $k = E(e^{\theta_t W_t}) = E(E(e^{\theta_t W_t}(\mathcal{F}_s)) \geq E(e^{\theta_t W_s})$, where the last inequality follows from the conditional Jensen's inequality (see [9]). Thus $\{\theta_t : t \geq 0\}$ is decreasing and $\lim_{t\to\infty} \theta_t = \theta^*$ exists. We need to show that $\theta^* > 0$. We will show that if $\{\theta_n\}$ is a subsequence, then $\lim_{n\to\infty} \theta_n = \theta^* > 0$. We begin by considering

$$E\left(e^{\theta_{n+1} W_{n+1}}\right) = E\left(e^{\theta_{n+1} W_n} E(e^{\theta_{n+1}(W_{n+1}-W_n)} | \mathcal{F}_n)\right) \tag{24}$$

Now, conditioned on $\mathcal{F}_n$,

$$W_{n+1} - W_n = e^{\alpha(n+1)} \left[\sum_{j=1}^{|Z_n|} \left(\sum_{i=1}^{|Z_1(x_i)|} V_{x_j}(y_i) - e^{-\alpha} V(x_j) \right) \right] \tag{25}$$

where $(x_1, \ldots, x_{|Z_n|})$ is the age chart at time n and $(y_1, y_2, \ldots, y_{|Z_{n+1}|})$ is the age chart at time $(n+1)$. Thus

$$E\left(e^{\theta_{n+1}(W_{n+1}-W_n)} | \mathcal{F}_n\right) = \prod_{j=1}^{|Z-n|} \Psi_{x_j}(\theta_{n+1} e^{\alpha(n+1)})$$

where $\Psi_{x_j}(\theta) = E\left(e^{\theta T(x_j)}\right)$ and

$$T(x_j) = \left[\sum_{i=1}^{|Z_1(x_j)|} V_{x_j}(y_i) \right] - e^{-\alpha} V(x_j)$$

Note that, since $E(T(x_j)) = 0$, $\sup_{y\geq 0} E(T^2(y)) < \infty$, and $\sup_{y} E(e^{\theta T(y)} T^2(y)) < \infty$ the following estimate prevails:

$$\Psi_{x_j}(\theta) \leq 1 + C\theta^2 \tag{26}$$

where C is a finite positive constant (see also [1] and [5]).

Using this estimate in (24) and the fact $\inf_{x\geq 0} V(x) > 0$ one has

$$E(e^{\theta_n W_n}) \leq E\left(e^{\theta_{n+1} W_n + C W_n \theta_{n+1}^2 e^{2\alpha(n+1)}}\right)$$

where C is a finite positive constant. From this it follows that

$$\theta_{n+1} \geq \frac{\theta_n}{1 + C\theta_{n+1 e^{2\alpha(n+1)}}}. \tag{27}$$

Iterating this inequality and letting $n \to \infty$ we get

$$\theta^* \geq \theta_1 \left\{ \prod_{j\geq 1} (1 + C\theta_j e^{2\alpha(j+1)} \right\}^{-1} > 0$$

concluding the proof of Theorem 8. □

Results of this type for single type branching processes were used by Biggins and Shanbhag [6] for studying certain infinite divisibility problems.

Theorem 9. *Under the assumptions of Theorem 8, there exist finite positive constants k and C such that*

$$P(|W_t - W| > \varepsilon) \leq Ce^{-ke^{-\alpha t/3}}.$$

Proof. Consider

$$P(W - W_t > \varepsilon) = E(P(W - W_t > \varepsilon | \mathcal{F}_t))$$

and observe that

$$W - W_t = \lim_{r\to\infty} (W_{t+r} - W_t).$$

Hence conditioned on $\mathcal{F}_t$,

$$W - W_t = e^{\alpha t} \sum_{j=1}^{|Z-t|} (W(x_j) - V(x_j))$$

where $W(x_j)$ is the limit of the martingale starting with a single particle of age x_j. Set $Y(x_j) = W(x_j) - V(x_j)$. Note that $\{Y(x_j)\colon j \geq 1\}$ is an independent sequence of mean 0 random variables. Thus

$$\begin{aligned} P(W - W_t > \varepsilon | \mathcal{F}_t) &\leq P\left(\frac{1}{|Z_t|} \sum_{j=1}^{|Z_t|} Y(x_j) > \frac{e^{\frac{-\alpha}{2}t}\varepsilon}{\sqrt{W_t}} \right) \\ &\leq E\left(e^{\frac{\theta}{\sqrt{|Z_t|}} \sum_{j=1}^{|Z_t|} Y(x_j)} \right) e^{\frac{-\theta e^{\frac{-\alpha t}{2}}\varepsilon}{\sqrt{W_t}}} \\ &\triangleq \left\{ \prod_{j=1}^{|Z_t|} \Psi_j \left(\frac{\theta}{\sqrt{|Z_t|}} \right) \right\} e^{\frac{-\theta e^{\frac{-\alpha t}{2}}\varepsilon}{\sqrt{W_t}}} \end{aligned}$$

Using an estimate similar to (25), we see that

$$\prod_{j=1}^{|Z_t|} \Psi_j \left(\frac{\theta}{\sqrt{|Z_t|}} \right) \leq \left(1 + \frac{\theta^2}{|Z_t|} \right)^{|Z_t|} \leq e^{\theta^2} \tag{28}$$

Thus

$$P(W - W_t > \varepsilon) \leq CE\left(e^{\frac{-\theta e^{-\alpha t/2}\varepsilon}{\sqrt{W_t}}} \right).$$

Now consider for $\lambda > 0$

$$\begin{aligned} E(e^{-\lambda \cdot \frac{1}{\sqrt{W_t}}}) &= \int_0^\infty e^{-\lambda u} P(\frac{1}{W_t} > u^2) du \\ &\leq C_1 \int_0^\infty e^{-\lambda u} e^{-\theta/u^2} du \text{ (by Theorem 8)} \\ &= C_1 \int_0^\infty e^{-u} e^{-\theta\lambda^2/u^2} du \end{aligned}$$

Consider

$$\begin{aligned} I(\lambda) = \int_0^\infty e^{-u} e^{-\theta\lambda^2/u^2} du &= \left(\int_0^{k(\lambda)} + \int_{k(\lambda)}^\infty \right) e^{-u} e^{-\frac{\theta\lambda^2}{u^2}} du \\ &\leq e^{-k(\lambda)} + e^{-\frac{\lambda^2}{k^2(\lambda)}} \end{aligned}$$

The value of $k(\lambda)$ which minimizes the RHS is $\lambda^{2/3}$ and hence

$$E(e^{-\lambda/\sqrt{W_t}}) \leq 2e^{-\lambda^{2/3}}$$

and thus

$$P(W - W_t > \varepsilon) \leq Ce^{-\theta^{2/3}e^{-\alpha t/3}\varepsilon^{2/3}}$$

A similar estimate prevails for $P(W_t - W > \varepsilon)$ thus completing the proof of Theorem 9. □

6. Open Problems. There are several unresolved questions that arise from our work. We list a few of them.

1. If G has a superexponentially decaying tail and $p_1 = 0$ we obtained the upper bound and lower bound for the decay rate of $\log F_t[0](s)$. Is it possible to improve this to a full convergence result as was done in [1] and [5]?
2. Theorem 1 is an analogue of Schroder equation for continuous time. What is the correct analogue for Böttcher case?
3. In [15] we obtained large deviation rates for the tails of W in a multitype branching process using Harris function and Karlin McGregor function. Is it possible to do this for the age-dependent case? One approach would be to study the Laplace transform of $W(x)$.
4. Is it possible to prove an analogue of Theorem 7 for the sub-exponential case by assuming only polynomial moments?

References.

1. Athreya, K. B. (1994): *Large deviation rates for branching processes - I, the single type case*, Annals of Applied Probability, to appear.
2. Athreya, K. B. and Kaplan (1976): *Convergence of the age distribution in the one-dimensional supercritical age dependent branching process*, Annals of Probability **4, No. 1**, 38-50.

3. Athreya, K. B. and Ney, P. E. (1972): *Branching Processes*, Springer-Verlag, New York.

4. Athreya, K. B. and Vidyashankar, A. N. (1993): *Large deviation results for branching processes, Stochastic Processes, a Festschrift in honor of Gopinath Kallianpur*, Springer-Verlag.

5. Athreya, K. B. and Vidyashnakar, A. N. (1994): *Large deviation rates for branching processes - II, the multitype case*, submitted.

6. Biggins, J. D. and Shanbhag, D. N. (1981): *Some divisibiity problems in branching processes*, Math. Proc. Camb. Phil. Soc. **90**, 321-330.

7. Dembo, Amir and Zeitouni, Ofer (1993): *Large deviations techniques and applications*, Jones and Bartlett Publishers, Boston.

8. Feller, N. (1966): *An introduction to probability theory and its applications*, vol. II, John Wiley & Sons, New York.

9. Durett (1991): *Probability: Theory and Examples*, Wadsworth and Brooks/Cole Advanced Books Software, Pacific Grove, CA.

10. Harris, T. E. (1963): *The theory of branching processes*, Springer-Verlag, New York.

11. Kuczek, Thomas (1982): *On the convergence of the empric age distribution for one dimensional super critical age dependent branching processes*, Annals of Probability **10, No. 1**, 252-258.

12. Nerman, O. (1981): *On the convergence of supercritical general (C-M-J) branching processes*, Z. Wahrsch. Verw. Gebiete **57**, 365-395.

13. Royden, H. L. (1987): *Real Analysis*, Macmillan Publishing Company, Third Edition.

14. Schuh, H-J. (1982): *Senata Constants for the supercritical Bellman-Harris Process*, Adv. Appl. Prob. **14**, 732-751.

15. Vidyashankar, A. N. (1994): *Large deviations for the tail behavior of W in a Multitype branching process*, submitted.

KESTEN-STIGUM TYPE THEOREM FOR BRWRE

A PAPER SUBMITTED TO JOURNAL OF
STOCHASTIC PROCESSES AND APPLICATIONS
ANAND N. VIDYASHANKAR

Departments of Mathematics & Statistics
Iowa State University, Ames, IA 50011

Abstract.

Let $\{Z^{(n)} : n \geq 0\}$ be a branching random walk in stationary ergodic environments starting with a single ancestor at the origin. Let $A_n(\theta) = \sum_{r=1}^{|Z^{(n)}|} e^{\theta z_{r,n}}$ and $P_n(\theta) = E(A_n(\theta))$. The sequence of random variables $\{W_n(\theta) = (P_n(\theta))^{-1} A_n(\theta), n \geq 1\}$ forms a non-negative martingale sequence and converges to $W(\theta)$ with probability 1 for almost all environments. In this paper we obtain necessary and sufficient conditions for the non-degeneracy of the limit.

1. Introduction.

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(H, \mathcal{H})$ be a measurable space of distributions of point processes on R (the real line). For each $n \geq 0$, let $\xi_n : (\Omega, \mathcal{H}, P) \to (H, \mathcal{H})$ be a stationary ergodic sequence of random variables and $\overline{\xi} = (\xi_0, \xi_1, \xi_2, \dots)$. In a branching random walk in random environments, an initial ancestor is born at the origin. He lives one unit of time and produces offspring. The positions of the offspring form a point process, centered at the position of the parent (which in this case is the origin), with distribution $\xi_0(\omega)$. The resulting population positions are $Z^{(1)}$. The population in the first generation reproduce independent of each other (and also independent of the previous generations) in exactly the same way as the initial ancestor, but this time the point process has distribution $\xi_1(\omega)$. The process continues and the resulting process $\{Z^{(n)}, n \geq 0\}$ is called the branching random walk in random environments. $\overline{\xi}$ is called the environmental sequence. Conditionally on the environment $\overline{\xi}$, $\{Z^{(n)}, n \geq 0\}$ is a time inhomogeneous branching random walk.

Some comments about the space $(H, \mathcal{H})$ are in order. We assume that the space H is topologized by weak convergence and since such a space can be metrized by Prohorov-Levy metric d, (H, d) is a metric space. It can be shown that (see [5]) this is a complete separable metric space. We also assume that the σ-field $\mathcal{H}$ is generated by the open sets produced by the metric d.

Let $H^\infty = H \times H \times H \times \ldots$ with the metric

$$\bar{d}(x, y) = \sum_{k \geq 0} \frac{1}{2^k} \frac{d(x_k, y_k)}{1 + d(x_k, y_k)}$$

where $x = (x_0, x_1, \ldots)$ and $y_k = (y_0, y_1, \ldots)$. Let $\mathcal{H}^\infty$ be the σ-algebra generated by the open sets. Define

$$T : (H^\infty, \mathcal{H}^\infty) \to (H^\infty, \mathcal{H}^\infty)$$

by

$$T(\overline{\xi}) = (\xi_1, \xi_2, \ldots)$$

and $\{T^{(k)}; k \geq 1\}$ to be the iterates of T. Also, for $n \geq 1$ let

$$\begin{aligned}
\mathcal{F}^{(0)} &= \sigma\langle \xi_0, \xi_1, \xi_2, \ldots \rangle \\
\mathcal{F}^{(n)} &= \sigma\langle Z^{(1)}, Z^{(2)}, \ldots, Z^{(n)}, \overline{\xi} \rangle
\end{aligned}$$

be the σ-algebras generated by the environmental sequence $\overline{\xi}$, and the environmental sequence and the process up to time n respectively. Let $|Z^{(n)}|$ denote the cardinality of $Z^{(n)}$. Then $|Z^{(n)}|$ is a Galton-Watson Process in stationary ergodic environments and is well studied in the literature (see [2], [6], and [7]).

We let $Z^{(n)} = \{z_{r,n} : r = 1, 2, \ldots |Z^{(n)}|\}$ and define for $\theta \in R$, $x \in R$, and $n \geq 1$, $A_n(\theta) = \sum_{r=1}^{|Z^{(n)}|} e^{\theta z_{r,n}} = \int_{-\infty}^{\infty} e^{\theta x} dZ^{(n)}(x)$ (We shall call this the A-transform set). $Z^{(n)}(x) = \#\{r : z_{r,n} \leq x\}$, $H_n(x) = E(Z^{(n)}(x)|\mathcal{F}^{(0)})$. Also let $F_{n-1}(x) = E(Z^{(n)}(x)|Z^{(n-1)} = \{0\}, \mathcal{F}^{(0)}), P_n(\theta) = E(A_n(\theta)|\mathcal{F}^{(0)}) = \int_{-\infty}^{\infty} e^{\theta x} H_n(dx) m_{\xi_{n-1}}(\theta) \triangleq m_{n-1}(\theta) = E(A_n(\theta)|Z^{(n)} = \{0\}, \mathcal{F}^{(0)}) = \int_{-\infty}^{\infty} e^{\theta x} F_n(dx)$, and $K(\theta) = E_{\xi_0}(\log m_0(\theta))$, where E_{ξ_0} denotes the expectation with respect to the distribution of ξ_0. Let $\mathcal{D}_K = \{\theta : E_{\xi_0}(|\log m_0(\theta)|) < \infty\}$. If we set $W_n(\theta) = (P_n(\theta))^{-1} A_n(\theta)$ for $\theta \in \mathcal{D}_K$, then one can show using the branching property (see [1]) that $\{(W_n(\theta), \mathcal{F}^{(n)}) : n \geq 1\}$ is a non-negative martingale sequence and hence converges with probability 1 to a random variable $W(\theta)$. The aim of this paper is to obtain necessary and sufficient conditions for the non-degeneracy of the limit when the generation size process, viz. $\{|Z^{(n)}| : n \geq 0\}$, is supercritical. Such a result for branching random walk in fixed environments was obtained by Biggins (see [3] and [4]).

The method of proof is an adaptation of the martingale truncation method of Tanny (see [7]) to the branching random walk case and it also provides a different proof of a similar result for branching random walk in fixed environments.

One of the motivations for this work comes from the study of "large deviations of $Z^{(n)}$". More precisely, when one considers the problem of determining the exact growth rate of $Z^{(n)}(nx)$, it turns out that $(E(Z^{(n)}(nx)))^{-1} Z^{(n)}(nx)$ converges, for a certain x, to the random-variable $W(\theta_x)$ (which is same as the martingale limit $W(\theta)$). The non-degeneracy of $W(\theta_x)$ implies that $(E(Z^{(n)}(nx)))^{-1}$ is the exact rate (see [8]).

The paper is organized as follows: section 2 briefly mentions the assumptions and section 3 is devoted to the main result of this paper.

2. Assumptions.

A1: $K(0) > 0$

A2: $\mathcal{D}_K = \{\theta : K(\theta) < \infty\}1$ is non-empty

Under A1, the branching process formed by the generation sizes is supercritical and hence the process survives with positive probability (see [2], [6]).

A3: $P(|\xi_n| \geq 1 \text{ for all } n \geq 1) = 1$, where ξ_n is the number of offsprings produced by a parent in the nth generation.

Under A3, the probability of extinction is 0. It should be noted that A3 entails no loss of generality since otherwise we need to condition on the set of non-extinction.

3. Main Result.

In this section we state and prove the main result. We start with the necessary condition. Before we embark into the first proposition, we introduce some more notation. We do this by setting $X_{n,s}(\theta) = \sum_{r=1}^{|T_{s,n-1}|} e^{\theta({}_s z_{r,n} - z_{s,n-1})}$; where ${}_s z_{r,n}$ are the positions of the offspring (who belong to the nth generation) coming from a parent at $z_{s,n-1}$ in the $(n-1)$th generation and $|T_{s,n-1}|$ is the number of offspring produced by that parent. We shall call this the X-transform of the parent at $z_{s,n-1}$. Using this notation one can express $A_n(\theta)$ as follows:

$$A_n(\theta) = \sum_{s=1}^{|z^{(n1)}|} e^{\theta z_{s,(n-1)}} X_{n,s}^{(\theta)}. \tag{1}$$

Also observe that conditionally on $\mathcal{F}^{(n-1)}, \{X_{n,s}(\theta) : 1 \leq s \leq |Z^{(n-1)}|\}$ are i.i.d. random variables. Let $G_n^\theta(x)$ denote their common distribution. We are now ready to state our first proposition.

Proposition 3.1. *Under A1-A3, a necessary condition for $W(\theta)$ to be non-degenerate is that*

$$\sum_{n\geq 1} \frac{1}{m_n(\theta)} \int_{P_{n+1}(\theta)}^{\infty} x dG_n^\theta(x) < \infty. \tag{2}$$

Proof. We start by defining a truncated version of $A_n(\theta)$. Set $A_0^t(\theta) = 1$ and $A_1^t(\theta) = X_{1,1}(\theta)I_{[0,CP_1(\theta)]}$. For $n \geq 2$ let $S_n^t = \{s : X_{n,s}(\theta) \leq CP_n(\theta)\}$

$$|S_n^t| = \text{ cardinality of } S_n^t$$

and

$$A_n^t(\theta) = \sum_{s \in S_n^t} e^{\theta z_{s,n-1}} X_{n,s}^t(\theta)$$

where

$$X_{n,s}^t(\theta) = X_{n,s}(\theta)I_{[0,CP_n(\theta)]}$$

and C is a finite positive constant.

The above quantities are related to a truncated branching random walk in varying environments $\{Z_t^{(n)}, n \geq 0\}$ constructed as follows. The process starts with one ancestor at the origin; the point process describing the positions of the first generation population is $Z^{(1)}$. If the transform of the ancestor is below his threshold (which is $CP_1(\theta)$), set $Z_t^{(1)} = Z^{(1)}$ (i.e. retain the entire first generation); else set $|Z_t^{(1)}| = 0$ (i.e. kill the entire first generation). This procedure is repeated for every individual in the first generation. Note that the threshold for the first generation population is $CP_2(\theta)$. The A-transform for this branching random walk is the same as $A_n^t(\theta)$ and the mean is $P_n^t(\theta) = \prod_{k=0}^{(n-1)} m_k^t(\theta)$ where

$$m_k^t(\theta) = \int x I_{[0,CP^{k+1}(\theta)]} dG_k^\theta(x)$$

for $k \geq 0$. Hence $\{(P_n^t(\theta)^{-1} A_n^t(\theta), \mathcal{F}^{(n)}) : n \geq 1\}$ is a non-negative martingale sequence. Let $W^t(\theta)$ be the limit. Hence $P(0 \leq W^t(\theta) < \infty | \mathcal{F}^{(0)}) = 1$.

Supposing one can show that given $\varepsilon > 0$, there exists a constant $C(\varepsilon, \theta, \bar{\xi})$ such that $P(A_n(\theta) = A_n^t(\theta) \text{ for all } n \geq 0 | \mathcal{F}^{(0)}) \geq 1 - \varepsilon$, then it would follow that on such a set for all $n \geq 0$,

$$W_n(\theta) = \frac{A_n(\theta)}{P_n(\theta)} = \frac{A_n^t(\theta)}{P_n^t(\theta)} \cdot \frac{P_n^t(\theta)}{P_n(\theta)}$$
$$= \left(\frac{A_n^t(\theta)}{P_n^t(\theta)}\right) \prod_{k=0}^{(n-1)} \frac{m_k^t(\theta)}{m_k(\theta)}$$
$$= \left(\frac{A_n^t(\theta)}{P_n^t(\theta)}\right) \prod_{k=0}^{n-1} \left(1 - \frac{1}{m_k(\theta)} \int_{CP_{k+1}(\theta)}^{\infty} x dG_k^\theta(x)\right)$$

and hence

$$W(\theta) = W^t(\theta) \left(\prod_{k \geq 0} \left(1 - \frac{1}{m_k(\theta)} \int_{CP_{k+1}(\theta)}^{\infty} x dG_k^\theta(x)\right)\right). \tag{3}$$

If (2) does not hold, then the RHS of (3) converges to 0 and hence $W(\theta)$ is 0 on a set of probability at least $1 - \varepsilon$; since ε is arbitrary $P(W(\theta) = 0) = 1$. Thus to conclude the proof, we need only establish that the aforesaid C can be chosen. To do this, note that

$$P\left(\sup_{n \geq 1} \frac{A_n(\theta)}{P_n(\theta)} < \infty | \mathcal{F}^{(0)}\right) = 1.$$

Hence given $\varepsilon > 0$, there exists $C(\varepsilon)$ such that

$$P\left(\sup_{n \geq 1} \frac{A_n(\theta)}{P_n(\theta)} \leq C | \mathcal{F}^{(0)}\right) \geq 1 - \varepsilon$$

which implies that

$$P\left(\sup_{n \geq 1} \sup_{1 \leq s \leq |Z^{(n-1)}|} (e^{\theta z_{s,n-1}} X_{n,s}(\theta)) \leq C P_n(\theta)\right) \geq 1 - \varepsilon$$

Hence $P(A_n(\theta) \neq A_n^t(\theta)$ for some $n \geq 1) \leq P(X_{n,s}(\theta) > C P_n(\theta)$ for some n and $s)$ which is bounded above by ε, thus completing the proof of Proposition 3.1.

We now turn our attention to the sufficient condition. In this case we shall see an interesting connection with the theory of large deviations. For this reason we shall strengthen some of our assumptions:

B1: Assume that $0 \in Int(\mathcal{D}_k)$ (i.e. 0 is an interior point of $\mathcal{D}_k$).

B2: $K(\theta)$ is a differentiable function of θ.

At this point we should remark that the assumption B_2 can be replaced by much weaker assumptions at the cost of adding new technicalities and we shall not pursue this line of approach. We state the next proposition whose proof is an elementary consequence of standard large deviation theory (see [3]).

Proposition 3.2. *Under A1, A3, B1, and B2*

$$\lim_{n\to\infty} \frac{1}{n} \log H_n(nx) = -I(x)$$

where

$$I(x) = \sup_\theta [\theta x - K(\theta)].$$

It can be seen from the large deviation theory that $\{x | I(x) < 0\}$ is an open interval (a, b) and for $x \notin (a, b)$, $I(x) \geq 0$. The next proposition is a simple consequence of the above observations and the Borel-Cantelli lemma.

Proposition 3.3. *For $x > b$, there exists $N_0(w, x, \bar{\xi})$ such that for all $n \geq N_0$*

$$Z^{(n)}[nx, \infty) = 0.$$

A similar result follows for $x < a$. These ideas motivate the following definition.

Definition. $SR = \{x | I(x) < 0\}$ is called the supercritical region of the BRWRE.

Observe that SR is non-empty since $0 \in SR$. We shall assume that

B3: For every $x \in SR$, there exists $\theta_x \in \mathcal{D}_k$ such that $K'(\theta_x) = x$.

From this point onwards we shall assume that such a θ_x has been chosen. Without loss of generality, we shall assume that $x = 0$. Under this situation $K(\theta) > 0$. We need one final assumption before we state the sufficient condition for non-degeneracy of $W(\theta)$.

B4: Assume that the environment $\bar{\xi}$ satisfies the following condition.

For every $Q > 0$, $\sum_{n\geq 0} \frac{1}{m_{2n}(\theta)} \int_{e^{nQ}}^{\infty} x dG_{2n}^{\theta}(x) < \infty$ a.e. on $E = \{\bar{\xi} : \sum_{n\geq 0} \frac{1}{m_n(\theta)} \int_{e^{nQ}}^{\infty} x dG_n^{\theta}(x) < \infty\}$. Our next lemma is technical and is needed in the proof of the sufficient condition.

Lemma 3.1. *Under B4, if* $\sum_{n\geq 0} \frac{1}{m_n(\theta)} \int_{e^{nQ}}^{\infty} x dG_n^{\theta}(x) < \infty$ *then* $\sum_{n\geq 0} \frac{1}{m_n(\theta)} \int_{e^{nQ/2}}^{\infty} x dG_n^{\theta}(x) < \infty$.

Proof.
$$\sum_{n\geq 0} \frac{1}{m_n(\theta)} \int_{e^{nQ/2}}^{\infty} x dG_n^{\theta}(x) = \left[\sum_{\substack{n\geq 0\\ n\,odd}} + \sum_{\substack{n\geq 0\\ n\,even}}\right] \left(\frac{1}{m_n(\theta)} \int_{e^{nQ/2}}^{\infty} x dG_n^{\theta}(x)\right)$$
$$\leq \sum_{n\geq 0} \frac{1}{m_{2n+1}(\theta)} \int_{e^{nQ}}^{\infty} x dG_{2n+1}^{\infty}(x) + \sum_{n\geq 0} \frac{1}{m_{2n}} \int_{e^{nQ}}^{\infty} x dG_{2n}^{\theta}(x)$$

Note that the first term is the same as the second term except that the environmental sequence is replaced by $T\bar{\xi}$. The lemma follows from B4.

Proposition 3.4. *Under A1, A3, and B1–B4 a sufficient condition for nondegeneracy of $W(\theta)$ is that (2) holds; viz.,*

$$\sum_{n\geq 1} \frac{1}{m_n(\theta)} \int_{P_{n+1}(\theta)}^{\infty} x dG_n^{\theta}(x) < \infty. \tag{2}$$

Proof. Since $K(\theta) > 0$ and $\frac{1}{n} \log P_n(\theta)$ converges to $K(\theta)$, given $\eta > 0$, there exists $N_0(\eta, \bar{\xi})$ such that $n(K(\theta) - \eta) \leq P_n(\theta) \leq n(K(\theta) + \eta)$ and hence (2) entails

$$\sum_{n\geq 0} \int_{e^{(n+1)(K(\theta)+\eta)}}^{\infty} x dG_n^{\theta}(x) < \infty \quad \text{with probability 1.}$$

We shall define another varying environment branching random walk similar to the one defined in Proposition 3.1 by replacing $CP_n(\theta)$ by $Dexp[\frac{(n+1)}{2}(K(\theta)+\eta)]$. Thus, as in that situation,

$$m_n^*(\theta) = \int_0^{De^{((n+1)(K(\theta)+\eta))/2}} x dG_n^\theta(x)$$

and

$$P_n^*(\theta) = \prod_{k=0}^{n-1} m_k^*(\theta).$$

We observe that there exists a D such that $m_n^*(\theta) > 0$ for all $n \geq 0$ and hence, (having chosen such a D) it follows that $\{(W_n^*(\theta), \mathcal{F}^{(n)}), n \geq 1\}$ where

$$W_n^*(\theta) = \frac{A_n^*(\theta)}{P_n^*(\theta)}$$

is a non-negative martingale sequence which converges to a non-negative random variable $W^*(\theta)$.

Suppose we can show that $\{W_n^*(\theta), n \geq 1\}$ is an L^2-bounded sequence; then it would follow that $E(W^*(\theta)|\mathcal{F}^{(0)}) = 1$. Hence

$$W_n(\theta) \geq W_n^*(\theta) \prod_{k=0}^{n-1} \frac{m_k^*(\theta)}{m_k(\theta)}$$

and letting $n \to \infty$

$$W(\theta) \geq W^*(\theta) \prod_{k \geq 0} \frac{m_k^*(\theta)}{m_k(\theta)} \text{ with probability 1.}$$

Consider

$$\prod_{k \geq 0} \frac{m_k^*(\theta)}{m_k(\theta)} = \prod_{k \geq 0} \left(1 - \frac{1}{m_k(\theta)} \int_{De^{((k+1)(K(\theta)+\eta)/2)}}^{\infty} x dG_k^\theta(x) \right). \tag{4}$$

Note that as D increases, the RHS of (4) increases to 1 by Lemma 3.4 and (2); thus

$$E(W(\theta)|\mathcal{F}^{(0)}) \geq E(W^*(\theta)|\mathcal{F}^{(0)}) = 1.$$

Also, by Fatou's lemma, $E(W(\theta)|\mathcal{F}^{(0)}) \leq 1$; whence $E(W(\theta)|\mathcal{F}^{(0)}) = 1$, completing the sufficient condition.

Thus to complete the proof of the theorem we need only establish that

$$\sup_{n\geq 0} Var(W_n^*(\theta)|\mathcal{F}^{(0)}) < \infty \quad \text{with probability 1.}$$

Consider

$$Var(W_n^*(\theta)|\mathcal{F}^{(0)}) = \frac{1}{(P_n^*(\theta))^2} Var(A_n^*(\theta))$$

But

$$\begin{aligned} Var(A_n^*(\theta)) &= E(A_n^*(\theta) - E(A_n^*(\theta)))^2 \\ &= E(A_n^*(\theta) - m_{n-1}^*(\theta)A_{n-1}^*(\theta) + m_{n-1}^*(\theta)A_{n-1}^*(\theta) - E(A_n^*(\theta)))^2 \\ &= E(A_n^*(\theta) - m_{n-1}^*(\theta)A_{n-1}^*(\theta))^2 + (m_{n-1}^*(\theta))^2 E(A_{n-1}^*(\theta) - E(A_{n-1}^*(\theta)))^2 \end{aligned}$$

Thus,

$$(5) \qquad Var(W_n^*(\theta)|\mathcal{F}^{(0)}) = E(Var(W_n^*(\theta)|\mathcal{F}^{(n-1)})|\mathcal{F}^{(0)}) + Var(W_{n-1}^*(\theta)|\mathcal{F}^{(0)}).$$

Iterating the above equality one obtains

$$(6) \qquad \sup_{n\geq 0} Var(W_n^*(0)|\mathcal{F}^{(0)}) \leq \sum_{n\geq 0} E(Var(W_n^*(0)|\mathcal{F}^{(n-1)})|\mathcal{F}^{(0)}).$$

But note that,

$$\begin{aligned} E(Var(W_n^*(\theta)|\mathcal{F}^{(n-1)})) &= \frac{1}{(P_n^*(\theta))^2} Var\left(\sum_{s\in S_{n-1}^t} e^{\theta z_{s,n-1}} X_{n,s}^*(\theta)|\mathcal{F}^{(n-1)}\right) \\ &\leq \frac{1}{(P_n^*(\theta))^2} \sum_{s\in S_{n-1}^t} e^{2\theta z_{s,n-1}} E(X_{n,s}^*(\theta))^2 \\ &= s_n^2(\theta)\frac{E(A_{n-1}^*(2\theta))}{(P_n^*(\theta))^2} \\ &= s_n^2(\theta)\frac{P_{n-1}^*(2\theta)}{(P_n^*(\theta))^2} \\ &= \left(\frac{P_{n-1}^*(2\theta)}{P_{n-1}^*(\theta)}\right)\frac{s_n^2(\theta)}{(P_{n-1}^*(\theta))(m_{n-1}^*(\theta))^2}. \end{aligned}$$

For large n, the first term is bounded and so

$$E(Var(W_n^*(\theta)|\mathcal{F}^{(n-1)})) \leq c(\theta, \bar{\xi}) \frac{s_n^2(\theta)}{P_n^*(\theta)(m_{n-1}^*(\theta))^2}.$$

Hence for large n,

$$\frac{s_n^2(\theta)}{P_n^*(\theta) m_{n-1}^*(\theta))^2} \leq C \frac{e^{(n+1)\frac{(k(\theta)+\eta)}{2}}}{P_n^*(\theta) m_{n-1}^*(\theta)} \tag{7}$$

where C is a finite positive constant depending on θ and the environment $\bar{\xi}$. The RHS of (7) can be rewritten as

$$c\frac{e^{D(n+1)\frac{(k(\theta)+\eta)}{2}}}{P_n(\theta) m_{n-1}(\theta)} \cdot \left(\frac{P_n(\theta)}{P_n^*(\theta)}\right)\left(\frac{m_{n-1}(\theta)}{m_{n-1}^*(\theta)}\right).$$

Since $\frac{P_n(\theta)}{P_n^*(\theta)}$ converges to $\prod\limits_{k\geq 1}\left(1 - \frac{1}{m_k(\theta)} \int\limits_{De^{(n+1)\frac{(k(\theta)+\eta)}{2}}}^{\infty} x dG_n^\theta(x)\right)$ by (2) and $\frac{m_{n-1}(\theta)}{m_{n-1}^*(\theta)}$ converges to 1 (also by (2)) we have that the RHS of (5) is $O\left(\frac{e^{(n+1)\frac{(k(\theta)+\eta)}{2}}}{P_n(\theta) m_{n-1}(\theta)}\right)$. First note that there exists $N_0(\bar{\xi}, \delta)$ such that for every $n \geq N_0$

$$e^{n(K(\theta)-\delta)} \leq \frac{1}{P_n(\theta)} \leq e^{n(K(\theta)+\delta)}$$

and $e^{-(n-1)\varepsilon} \leq \frac{1}{m_{n-1}(\theta)} \leq e^{(n-1)\varepsilon}$ and hence

$$\frac{e^{D(n+1)(k(\theta)+\eta)}}{P_n(\theta) m_{n-1}(\theta)} \leq C e^{n(\frac{k(\theta)+\eta}{2} - k(\theta)+\delta+\varepsilon)}.$$

Choosing $\left(\frac{\eta}{2} + \delta + \varepsilon\right) < \frac{K(\theta)}{2}$, we have that $\sum\limits_{n\geq 10} E(Var(W_n^*(\theta)|\mathcal{F}^{(n-1)})|\mathcal{F}^{(0)}) < \infty$, thus completing the proof of our proposition.

Combining Propositions 3.1 and 3.4 we have

Theorem 1. *Under A1–A3 and B1–B4, a necessary and sufficient condition for* $E(W(\theta)|\mathcal{F}^{(0)}) = 1$ *is that*

$$\sum_{n\geq 0} \frac{1}{m_n(\theta)} \int\limits_{P_{n+1(\theta)}}^{\infty} x dG_n^\theta(x) < \infty. \tag{2}$$

We now focus on producing conditions equivalent to (2). We shall also show that these conditions reduce to a familiar "$z \log z$" condition used in branching theory.

Theorem 2. *Under A1–A3 and B1–B4, a necessary and sufficient condition for $E(W(\theta)|\mathcal{F}^{(0)}) = 1$ is that*

$$\sum_{n\geq 0} \frac{1}{m_n(0)} \int_{e^{n\beta}}^{\infty} x dG_n^\theta(x) < \infty \tag{8}$$

for some $\beta > 0$. Furthermore (5) holds for some β if and only if it holds for all $\beta > 0$.

Proof. From Theorem 1, it is sufficient to show that if (5) holds for some $\beta > 0$ then it holds for all $\beta > 0$. Assume that (5) holds for $\beta = \beta_0$. Then $\sum_{n\geq 0} \frac{1}{m_n(\theta)} \int_{e^{n\beta_0/2}}^{\infty} x dG_n^\theta(x) <$ $\sum_{n\geq 0} \frac{1}{m_{2n}(\theta)} \int_{c^{n\beta_0}} x dG_{2n}^\theta(x) + \sum_{n\geq 0} \frac{1}{m_{2n+1}(\theta)} \int_{e^{n\beta_0}}^{\infty} dG_{2n+1}^\theta(x) < \infty$.
One can repeat this argument to conclude that 5 holds for all $\beta > 0$. □

Our next theorem specializes to the case of i.i.d. environments. Note that in this case the condition B4 is automatically satisfied.

Theorem 3. *Let $\{Z^{(n)} : n \geq 0\}$ be a BRWRE satisfying A1, A3 and B1–B3. Then a necessary and sufficient condition for $E(W(\theta)|\mathcal{F}^{(0)}) = 1$ is that* $E_{\xi_0}\left(E\left(\frac{A_1(\theta)\log^+ A_1(\theta)}{m_0(\theta)}\right)\right) < \infty$.

Proof. To prove the theorem, all we need to do is to show that (8) is equivalent to $E_{\xi_0}\left(E\left(\frac{A_1(\theta)\log^+ A_1(\theta)}{m_0(\theta)}\right)\right) < \infty$.

Consider the sequence $\{\frac{1}{m_n(\theta)} \int_{e^{n\beta}}^{\infty} x dG_n^\theta(x) : n \geq 0\}$ of i.i.d. random variables which are bounded by 1. By (8) and Kolmogorov's three series theorem,

$$E(\sum_{n\geq 0} \frac{1}{m_n(\theta)} \int_{e^{n\beta}}^{\infty} x dG_n^\theta(x)) < \infty. \tag{9}$$

By monotone convergence theorem, (9) is the same as

(10)

$$\sum_{n\geq 0} E_{\xi_0}\left(\frac{1}{m_0(0)} \int_{e^{n\beta}}^{\infty} xd\}_0^\theta(x)\right) = E_{\xi_0}\left(\frac{1}{m_0)\theta)} \int_0^{\infty} x \left(\sum_{n\geq 0} I_{[e^{n\beta},\infty)}(x)\right) dG_0^\theta(x)\right).$$

But the estimate (for $\beta > 0$ and $x \geq 1$)

$$\frac{1}{\beta}\log x \leq \sum_{n\geq 0} I_{[e^{n\beta},\infty)}(x) \leq \frac{1}{\beta}\log x + 1 \tag{11}$$

implies that the right hand side of (10) is bounded above and below by $KE_{\xi_0}\left(E\left(\frac{A_1(\theta)\log^+ A_1(\theta)}{m_0(\theta)}\right)\right)$ for some constant K, thus completing the proof of the theorem. □

The above theorem in particular gives a necessary and sufficient condition for the non-degeneracy of $W(\theta)$ in fixed environments, a result first proved in Biggins [3]. Finally, we also have (using the above line of arguments) that in the stationary ergodic case that "$z\log z$" is a sufficient condition for the non-degeneracy. We shall state it as a theorem and omit the proof.

Theorem 4. *Under A1–A3 and B1–B4 a sufficient condition for $E(W(\theta)|\mathcal{F}^{(0)}) = 1$ is that*

$$E_{\xi_0}\left(E\left(\frac{A_1(\theta)\log^+ A_1(\theta)}{m_0(\theta)}\right)\right) < \infty.$$

Finally we shall state an unresolved problem which naturally arises from the results proved here.

Open Problem: The condition on the environmental sequence is needed only for a technical calculation. Is it possible to remove this hypothesis?

References.

1. Athreya, K.B. and Kaplan, N. (1978): *Additive property and its applications to branching processes*, Branching Processes, ed. A. Joffe and P.E. Ney, 27–60.
2. Athreya, K.B. and Karlin, S. (1971): *On branching processes with random environments I*: Extinction probabilities, Ann. Math. Stat., 42, 1499–1520.
3. Biggins, J.D. (1977): *Martingale convergence in the branching random walk*, J. Appl. Prob., 14, 25–37.

4. Biggins, J.D. (1989): *Uniform convergence of martingales in the one-dimensional branching random walk*, In selected proceedings of the symposium on Applied Probability, Sheffield (I. V. Basawa and R. L. Taylor, eds.), vol. 18, IMS, Hayward, California, pp. 159-173.

5. Kallenberg (1983): *Random Measures*, Akademie–Verlag, Berlin, and Academic Press, New York.

6. Tanny, D. (1977): *Limit theorems for branching processes in random environment*, Annals of Probability, 5, 100–116.

7. Tanny, D (1988): *A necessary and sufficient condition for a branching process in a random environment to grow like the product of its means*, Stochastic Processes and their Applications, 28, 123–139.

8. Vidyashankar, A.N. (1994): *Growth rates for branching random walk in random environments*, submitted.

GROWTH RATES FOR BRANCHING RANDOM WALK IN RANDOM ENVIRONMENTS (BRWRE)

A PAPER SUBMITTED TO ANNALS OF APPLIED PROBABILITY

ANAND N. VIDYASHANKAR

Departments of Mathematics and Statistics
Iowa State University
Ames, IA 50011

Abstract. Let $\{Z^{(n)}: n \geq 0\}$ denote a branching random walk in stationary ergodic environments. This paper is concerned with the rate of growth of $Z^{(n)}(nx)$, the number of particles living to the left of nx. It is shown that , under suitable moment and regular conditions (for a certain x), $Z^{(n)}(nx)$ normalized by its expectation converges to a random variable $W(\theta_x)$.

1. Introduction. Let $(\Omega, \mathcal{F}, P)$ be a probability space and H denote the space of all distributions of point processes on R (the real line) with the topology of weak convergence. It is known that this topology is metrizable by the Prohorov metric d (see [7]) and the space (H, d) is a complete separable metric space. Let $\mathcal{H}$ be the σ-field generated by the open sets. For every $n \geq 0$, let $\xi_n: (\Omega, \mathcal{F}, P) \to (H, \mathcal{H})$ be a sequence of stationary ergodic random variables. Further, let $\overline{\xi} = (\xi_0, \xi_1, \xi_2, \dots)$. A branching random walk in random environments starts with one ancestor at the origin. He lives one unit of time and produces offspring, whose positions are centered at the position of their parent and forms a point process $Z^{(1)}$ with distribution $\xi_0(w)$. Each of the first generation parents in turn live one unit of time and are replaced by their progeny in exactly the same way as the initial ancestor, but the point process of positions have distribution ξ_1. Call the positions of the second generation population $Z^{(2)}$. The process continues and the resulting point process of positions$\{Z^{(n)}: n \geq 0\}$ is called the branching random walk in random environments. The sequence $\overline{\xi} = (\xi_0, \xi_1, \xi_2, \dots)$ is called the environmental sequence.

Our aim in this paper is to study the process $Z^{(n)}(nx)$, the number of particles of the nth generation living to the left of nx. In fact, we shall show, under further moment and regularity conditions, that $Z^{(n)}(nx)$, normalized by $E(Z^{(n)}(nx)|\overline{\xi})$ converges to a random variable $W(\theta_x)$ with probability 1. This random variable turns out to be the same as the random variable that comes up as a limit of a martingale sequence in BRWRE. In fact if $x = \infty$, $\{(E(Z^{(n)}(nx))^{-1}Z^{(n)}(nx)n \geq 1\}$ is a martingale sequence and its limit is well studied in the literature (see [3], [4], [9], and [10]). In a separate paper (see [11]) we have investigated the necessary and sufficient conditions for the non-degeneracy of the martingale $W(\theta_\ell)$. The moment conditions that we impose here will be more than sufficient to guarantee the non-triviality of the limit $W(\theta_x)$.

Results of this kind for fixed environments was first obtained by Biggins ([4], and [5]). Similar results for multi-type branching random walk was obtained by Bramson, Ney & Tao (see [6]). To prove our main theorem we need a local limit theorem for sums of independent random variables. We shall discuss this in Section 3 and in Section 4 we shall discuss the main theorem. Section 2 will be devoted to notations, assumptions and some preliminary results.

2. Notations, Assumptions and Preliminary Results.

Notations.

N1 $Z^{(n)}(x)$ is the number of particles of the nth generation living to the left of x.

N2 $Z^{(n)}[x, \infty)$ is the number of particles of the nth generation living to the right of x.

N3 $|Z^{(n)}|$ is the number of particles in the nth generation.

N4 $Z^{(n)} = \{z_{r,n} : r = 1, 2, \ldots, |Z^{(n)}|\}$ is the point process describing the positions of the nth generation population.

N5 $A_n(\theta) = \sum_{r=1}^{|Z^{(n)}|} e^{\theta z_{r,n}}$

N6 $\mathcal{F}^{(0)} = \sigma\langle \xi_0, \xi_1, \ldots \rangle$

N7 For $n \geq 1$, $\mathcal{F}^{(n)} = \sigma < Z^{(0)}, Z^{(1)}, \ldots, Z^{(n)}, \bar{\xi}\rangle$

N8 For $n \geq 1$, $F_{n-1}(x, \xi) = E(Z^{(n)}(x)|Z^{(n-1)} = \{0\}, \xi)$

N9 For $n \geq 0$, $H_n(x, \xi) = E(Z^{(n)}(x)|Z^{(0)} = \{0\}, \xi)$ using branching property one can show that $H_n(x, \xi) = F_{n-1} * F_{n-2} * \ldots F_0(x, \xi)$

N10 For $n \geq 0$, $m_n(\theta) = \int_{-\infty}^{\infty} e^{\theta x} dF_n(x) = E(A_n(\theta)|Z^{(n-1)} = (0), \xi)$

N11 For $n \geq 0$, $P_n(\theta, \xi) = \int_{-\infty}^{\infty} e^{\theta x} dH_n(x) = E(A_n(\theta)|Z^{(0)} = \{0\}, \xi) = \prod_{k=0}^{(n-1)} m_k(\theta)$, and for $k \geq 0$ $P_n(\theta, T^\ell \xi) = \prod_{k=\ell}^{(n+\ell-1)} m_k(\theta)$.

N12 For $n \geq 0$, $\overline{F}_n(x, \xi) = \frac{e^{\theta x} F_n(x,\xi)}{m_n(\theta)}$ and $\overline{H}_n(x, \xi) = \frac{e^{\theta x} H_n(x,\xi)}{P_n(\theta,\xi)}$.

N13 For $n \geq 0$, $\sigma_n^2 = \int_{-\infty}^{\infty} x^2 d\overline{F}_n(x, \xi)$ and $\gamma_n = \int_{-\infty}^{\infty} x^3 d\overline{F}_n(x, \xi)$, and $\nu_n(t) = \int_{-\infty}^{\infty} e^{itx} d\overline{F}_n(x, \xi)$.

N14 $B_n = \sum_{j=0}^{(n-1)} \sigma_j^2$.

N15 $K(\theta) = E_{\xi_0}(\log m_0(\theta))$ where E_{ξ_0} is the expectation with respect to the distribution of the random variable ξ_0.

N16 $\mathcal{D}_k = \{\theta : K(\theta) < \infty\}$.

N17 For $x \in R, I(x) = \sup_{\theta \in R}\{\theta x - K(\theta)\}$ and $SR = \{x | I(x) < 0\}$.

Assumptions.

A1 $E_{\xi_0}(\log m_0(0)) > 0$

A2 $\mathcal{D}_k$ is non-empty and $0 \in \text{Int}(\mathcal{D}_k)$ where $\text{Int}(\mathcal{D}_k)$ is the interior of $\mathcal{D}_k$. It can be seen using standard large deviation theory ideas that SR is an open interval (a, b).

A3 For every $x \in (a, b)$, there exists $\theta_x \in \mathcal{D}_k$ such that

$$K'(\theta_x) = x$$

The region SR is called the supercritical region of the BRWRE (see [11]). A1 implies that the BPRE formed by the generation sizes survives with positive probability (see [2]). Hence the next assumption entails no loss of generality.

A4 For almost all environments, the probability of producing zero offspring is 0.

A5 $E_{\xi_0}(\gamma_0) < \infty, \quad E_{\xi_0}(\sigma_0^2) > 0.$

A6 $\int\limits_{|t|>\varepsilon} \left(\prod\limits_{k=0}^{(n-1)} |\nu_k(t)| \right) dt = \mathbf{O}\left((B_n)^{-\frac{1}{2}} \right).$

A7 $E_{\xi_0}\left(E\left(\frac{A_1(\theta) \log_+^{\frac{5}{2}} A(\theta)}{M_0(\theta)} \right) \right) < \infty.$

A8 $\inf\limits_{k\geq 0} m_k(\theta) > 1$ for almost all environments

A9 $m_k' = 0$, for all $k \geq 0$ for almost all environments.

A10 Assume that the environment $\bar{\xi}$ satisfies the following condition: $\sum\limits_{n\geq 0} \frac{1}{m_{2n}(\theta)} \int\limits_{e^{nQ}}^{\infty} x dG_{2n}^{\theta}(x) < \infty$ a.e. on $E = \{\bar{\xi} : \sum\limits_{n\geq 0} \frac{1}{m_n(\theta)} \int\limits_{e^{nQ}}^{\infty} x dG_n^{\theta}(x) < \infty\}$ where $Q > 0$ and G_n^{θ} is the common conditional distribution of $X_{n,s}(\theta)$ which is the same as $\sum\limits_{r=1}^{|T_{s,n-1}|} e^{\theta({}_s z_{r,n} - z_{s,n-1})}$ where ${}_s z_{r,n}$ are the positions of the offspring (who belong to the nth generation) coming from a parent at $z_{s,n-1}$ in the $(n-1)$th generation and $|T_{s,n-1}|$ is the number of offspring produced by that parent.

3. Local Limit Theorem. In this section we shall prove a local limit theorem for sums of independent random variables. Our method of proof uses characteristic functions and is akin to the method adopted by Petrov (see [8]). Let $\{X_n : n \geq 1\}$ be a sequence of independent random variables with $EX_j = 0$, and $E|X_j^3| <$

be a sequence of independent random variables with $EX_j = 0$, and $E|X_j^3| < \infty$ for all $n \geq 1$. Set $\tilde{\sigma}_j^2 = EX_j^2, \tilde{B}_m^n = \sum_{j=m}^{n} \tilde{\sigma}_j^2, S_m^n = \sum_{j=m}^{n} X_j, \nu_n(t) = E\left(e^{itX_n}\right)$, $F_m^n(x) = P\left[(\tilde{B}_m^n)^{-\frac{1}{2}} S_m^n \leq x\right]$ and $L_m^n = (\tilde{B}_m^n)^{-\frac{3}{2}} \sum_{j=m}^{n} E(|X_j|^3)$. If $m = 0$, we shall denote B_0^n by B_n, p_0^n by p_n and S_0^n by S_n. Let $f_n(t)$ denote the characteristic function of the random variable $(\tilde{B}_n)^{-\frac{1}{2}} S_n$. The following estimate on the characteristic function is from Petrov ([8], page 109).

$$|f_n(t) - e^{-\frac{t^2}{2}}| \leq 16 L_n |t|^3 e^{-\frac{t^2}{3}} \text{ for } |t| \leq (4L_n)^{-1} \tag{1}$$

We are now ready to state our main theorem of this section.

Theorem 1. *Let $\{X_n : n \geq 1\}$ be a sequence of independent random variables with $EX_n = 0$ and $EX_n^2 < \infty$ for all $n \geq 1$. Assume that $\tilde{B}_m^n \to \infty$ as m and $n \to \infty$ such that $(n - m) \to \infty$. Further, let $\sum_{j=m}^{n} E(X_j^3) = \mathbf{O}(\tilde{B}_m^n)$ and for every $\varepsilon > 0$ $\int_{|t|>\varepsilon} \prod_{j=m}^{n} |\nu_j(t)dt = \mathbf{O}((\tilde{B}_m^n)^{-1})$. Then for sufficiently large n and m such that $(n - m)$ is large, there exists an everywhere continuous derivative p_m^n of F_m^n such that*

$$\sup_x |p_m^n(x) - \phi(x)| = \mathbf{O}\left((\tilde{B}_m^n)^{-1}\right)$$

where $\phi(x)$ is the standard normal density.

Proof. Let $f_m^n(t)$ denote the characteristic function of F_m^n. To establish the theorem we shall first show that if n and m are large such that $(n - m)$ is large that $f_m^n(\cdot)$ is absolutely Riemann integrable. Consider

$$\begin{aligned}
\int_{-\infty}^{\infty} |f_m^n(t)| dt &\leq \int_{-\infty}^{\infty} |f_m^n(t) - e^{-\frac{t^2}{2}}| dt + \sqrt{2\pi} \\
&\leq \int_{|t|>(4L_m^n)^{-1}} |f_m^n(t) - e^{-\frac{t^2}{2}}| dt + \int_{|t|\leq(4L_m^n)^{-1}} |f_m^n(t) - e^{-\frac{t^2}{2}}| dt + \sqrt{2\pi}
\end{aligned}$$

Set $I_1 = \int\limits_{|t| \leq (4L^n m)^{-1}} |f^n_m(t) - e^{-\frac{t^2}{2}}|dt$. Also observe that

$$\begin{aligned}
\int\limits_{|t|\geq (4L^n_m)^{-1}} |f_n m(t) - e^{-\frac{t^2}{2}}|dt &\leq \int\limits_{|t|\geq (4L^n_m)^{-1}} |f^n_m(t)|dt + \int\limits_{|t|\geq (4L^n_m)^{-1}} e^{-\frac{t^2}{2}} dt \\
&= \qquad I_2 \qquad\qquad + \qquad I_3
\end{aligned}$$

Note that

$$\begin{aligned}
I_2 &\leq \int\limits_{|t|\geq (4L^n_m)^{-1}} |f^n_m(t) dt \\
&= \int\limits_{|t|\geq (4L^n_m)^{-1}} \prod_{j=m}^{n} \nu_j(\frac{t}{\sqrt{\tilde{B}^n_m}}) dt \\
&\leq \sqrt{\tilde{B}^n_m} \int\limits_{|t|\geq (4C)^{-1}} \left[\prod_{j=m}^{n} \nu_j(t) \right] dt \quad (\text{since } L^n_m = \mathbf{O}\left((\tilde{B}^n_m)^{-\frac{1}{2}} \right)) \\
&= \mathbf{O}\left((\tilde{B}^n_m)^{-\frac{1}{2}} \right)
\end{aligned}$$

and $\quad I_3 \leq 8L^n_m e^{-\frac{1}{32(L^n_m)^2}}$.

Finally, from (1)

$$\int\limits_{|t|\leq (4L^n_m)^{-1}} |f^n_m(t) - e^{-\frac{t^2}{2}}|dt \leq 16L^n_m \int\limits_{|t|\leq (4L^n_m)^{-1}} |t|^3 e^{-\frac{t^2}{3}} dt.$$

Plugging in these estimates, the integrability of $f^n_m(\cdot)$ follows. Finally,

$$\begin{aligned}
\sup_x |p^n_m(x) - \phi(x)| &\leq \left(\sqrt{2\pi}\right)^{-1} \int_{-\infty}^{\infty} |f^n_m(t) - e^{-\frac{t^2}{2}}|dt \\
&\leq I_1 + I_2 + I_3 = \mathbf{O}\left((\tilde{B}^n_m)^{-\frac{1}{2}} \right)
\end{aligned}$$

thus completing the proof of Theorem 1. □

Corollary 3.1. *Under the conditions of Theorem 1, for any directly Riemann integrable function f*

$$\sup_y \left| \int_{-\infty}^{\infty} f(x+y)p_n\left(\frac{x}{\sqrt{\tilde{B}_n}}\right) dx - \int_{-\infty}^{\infty} f(x)\phi\left(\frac{x-y}{\sqrt{\tilde{B}_n}}\right) dx \right|$$

converges to 0 *as* $n \to \infty$.

Note that $\overline{F}_n(\cdot,\xi)$ *and* $\overline{H}_n(\cdot,\xi)$ *are distribution functions even through* F_n *and* H_n *are not. Our next proposition specializes Corollary 3.1 to the sequence* $\overline{H}_n$.

Proposition 3.1. *Under A5 and A6, there exists* $N(\overline{\xi})$ *such that for all* $n \geq N(\overline{\xi})$, $\overline{F}_n(x)$ *is differentiable with* $\overline{F}'_n(x) = \overline{p}_n(x)$ *and*

(i) $\sup_x |\overline{p}_n(x) - \phi(x)| = \mathbf{O}(B_n^{-\frac{1}{2}})$ *with probability 1 and*

(ii) *for any absolutely integrable function f*

$$\sup_y \left| \sqrt{B_n} \int_{-\infty}^{\infty} f(x+y)d\overline{H}_n(x) - \int_{-\infty}^{\infty} f(x)\phi\left(\frac{x-y}{\sqrt{B_n}}\right) dx \right|$$

converges to 0 *w.p.1.*

Proof. Since $E(\sigma_0^2) > 0, B_n \to \infty$ by the ergodic theorem. Another application of ergodic theorem also yields

$$\sum_{k=0}^{n-1} \gamma_k = \mathbf{O}(B_n).$$

The proposition follws from Theorem 1 and Corollary 3.1.

4. Growth rates for $Z^{(n)}$. In this section, we seek to prove an analogue of the second statement of the Proposition 3.1. We shall now state the main theorem of this paper.

Theorem 2. *Under A1 - A10, for almost all environments*

$$\frac{\sqrt{2\pi B_n(\xi)}}{P_n(\theta)} \int_{-\infty}^{\infty} e^{\theta x} f(x) dZ^{(n)}(x) \to W(\theta) \int_{-\infty}^{\infty} e^{\theta x} f(x) dx$$

with probability 1, where f is such that $\int_{-\infty}^{\infty} e^{\theta x} f(x)dx < \infty$.

Before we proceed with the proof of Theorem 2 we observe, by setting $\overline{Z}^{(n)}(\cdot) = \frac{e^{\theta x} Z^{(n)}}{P_n(\theta,\xi)}$, *that the statement of Theorem 2 is equivalent to*

$$T_n \triangleq \sqrt{2\pi B_n(\xi)} \int_{-\infty}^{\infty} f(x) d\overline{Z}^{(n)}(x) \to W(\theta) \int_{-\infty}^{\infty} f(x)dx$$

with probability 1 for almost all environments.

Our method of proof of Theorem 2 is to show that T_n is close to its conditional expected value for a suitably chosen time $l(n) < n$ and then to show that the conditional expectation converges to the desired limit. This idea goes back to Asmussen and Kaplan (see [1]) and has been successfully employed in this context by Biggins [4] and Bramson, et al [6]. Thus we are looking at the decomposition

$$T_n = \left(T_n - E(T_n|\mathcal{F}^{(\ell)})\right) + E\left(T_n|\mathcal{F}^{(\ell)}\right). \tag{2}$$

We shall choose for $0 < \alpha < 1$

$$\ell(n) = [j^\alpha] \text{ where } j^3 \le n < (j+1)^3 \tag{3}$$

where j is an integer. This choice of ℓ was employed by Biggins [4]. Our next proposition establishes the convergence of $E(T_n|\mathcal{F}^{(\ell)})$ to $W(\theta) \int_{-\infty}^{\infty} f(x)dx$.

Proposition 4.1. *$E(T_n|\mathcal{F}^{(\ell)})$ converges to $W(\theta) \int_{-\infty}^{\infty} f(x)dx$ as $n \to \infty$ with probability 1 for almost all environments.*

Proof.

$$E(T_n|\mathcal{F}^{(\ell)}) = \sqrt{2\pi B_n} E\left[\int_{-\infty}^{\infty} f(x) d\overline{Z}^{(n)}(x) | \mathcal{F}^{(\ell)}\right].$$

Consider

$$E\left[\int_{-\infty}^{\infty} f(x)d\overline{Z}^{(n)}(x)|\mathcal{F}^{(\ell)}\right] = E\left[\sum_{r=1}^{|Z^{(n)}|} \frac{f(z_{r,n})e^{\theta z_{r,n}}}{P_n(\theta,\xi)}|\mathcal{F}^{(\ell)}\right]$$

$$= \frac{1}{P_n(\theta,\xi)} \sum_{s=1}^{|Z^{(\ell)}|} e^{\theta z_{s,l}} E\left[\sum_{r\in S_{s,\ell}} f({}_s z_{r,n}) e^{\theta({}_s z_{r,n} - z_{s,l})}\right]$$

where $S_{s,\ell} = \{r\colon z_{r,n}$ is the position of a offspring in the nth generation whose parent was at $z_{s,l}$ in the the ℓth generation $\}$. But

$$E\left[\sum_{r\in S_{s,\ell}} f({}_s z_{r,n}) e^{\theta({}_s z_{r,n} - z_{s,\ell})}|\mathcal{F}^{(\ell)}\right] = \left[\prod_{k=\ell}^{n-1} m_k(\theta)\right] E\left[\int_{-\infty}^{\infty} f(x+z_{s,\ell})d\overline{Z}^{(n-\ell)}(x, T^\ell\xi)|\mathcal{F}^{(\ell)}\right]$$

$$= \left[\prod_{k=\ell}^{(n-1)} m_k(\theta)\right]\left[\int_{-\infty}^{\infty} f(x+z_{s,\ell})d\overline{H}_{n-\ell}(x, T^\ell\xi)\right]$$

and so

$$E\left(T_n|\mathcal{F}^{(\ell)}\right) = \frac{\sqrt{2\pi B_n(\xi)}}{P_\ell(\theta,\xi)} \sum_{s=1}^{|z^{(\ell)}|} e^{\theta z_{s,\ell}} \int_{-\infty}^{\infty} f(x+z_{s,\ell})d\overline{H}_{n-\ell}(x) \tag{4}$$

Now the RHS of the above equation is bounded above by (for large n)

$$\Bigg[k\sum_{s=1}^{|z^{(\ell)}|} e^{\theta z_{s,\ell}}\left(\int_{-\infty}^{\infty} f(x+z_{s,\ell})d\overline{H}_{n-\ell}(x) - \frac{1}{\sqrt{B_{n-\ell}(T^\ell\xi)}}\int_{-\infty}^{\infty} f(x)\phi\left(\frac{x-z_{s,\ell}}{\sqrt{B_{n-\ell}(T^\ell\xi)}}\right)dx\right)$$

$$+\frac{1}{P_\ell(\theta,\xi)}\sum_{s=1}^{|Z^{(\ell)}|} e^{\theta z_{s,\ell}} \int_{-\infty}^{\infty} f(x)\left[\phi\left(\frac{x-z_{s,\ell}}{\sqrt{B_{n-\ell}(T^\ell\xi)}}\right) - \phi\left(\frac{x}{\sqrt{B_{n-\ell}(T^\ell\xi)}}\right)\right]dx$$

$$+\frac{1}{P_n(\theta,\xi)}\sum_{s=1}^{|Z^{(\ell)}|} e^{\theta z_{s,\ell}} \int_{-\infty}^{\infty} f(x)\phi\left(\frac{x}{\sqrt{B_{n-\ell}(T^\ell\xi)}}\right)dx\Bigg] \text{ for some constant } k.$$

The last term inside the paranthesis converges to $W(\theta)\int_{-\infty}^{\infty} f(x)dx$. The first term inside the paranthesis converges to 0 by Theorem 3.1. We shall now show that the middle term converges to 0. Consider

$$|\frac{1}{P_\ell(\theta)}\sum_{s=1}^{|Z^{(\ell)}|} e^{\theta z_{s,\ell}} \int_{-\infty}^{\infty} f(x)\left(\phi\left(\frac{x - z_{x,\ell}}{\sqrt{B_{n-\ell}(T^\ell\xi)}}\right) - \phi\left(\frac{x}{\sqrt{B_{n-\ell}(T^\ell\xi)}}\right)\right) dx|$$

$$\leq \frac{K}{P_\ell(\theta,\xi)}\sum_{s=1}^{|Z^{(\ell)}|} e^{\theta z_{s,\ell}} \frac{|z_{s,\ell}|}{\sqrt{B_{n-\ell}(T^\ell\xi)}} = K\frac{\sqrt{B_n(\xi)}}{B_{n-\ell}(T^\ell\xi}\left(\frac{1}{P_\ell(\theta,\xi)}\sum_{s=1}^{|Z^{(\ell)}|}\frac{e^{\theta z_{s,\ell}}|z_{s,\ell}|}{\sqrt{B_n(\xi)}}\right)$$

By ergodic theorem, and the assumption that $E(\sigma_0^2) > 0$, there exists $N_0(\overline{\xi})$ such that for all $n \geq N_0$

$$\sqrt{n(E(\sigma_0^2) - \delta)} \leq \sqrt{B_n} \leq \sqrt{n(E(\sigma_0^2) + \delta)}.$$

Thus the quantity inside the paranthesis is bounded above by

$$\frac{K}{\sqrt{n}}\frac{1}{P_\ell(\theta,\xi)}\sum_{s=1}^{|Z^{(\ell)}|} e^{\theta z_{s,\ell}}|z_{s,\ell}| \tag{5}$$

which in turn is bounded above by $\frac{K}{j^{\frac{3}{2}}}\frac{1}{P_\ell(\theta,\xi)}\sum_{s=1}^{|Z^{(\ell)}|} e^{\theta z_{s,\ell}}|z_{s,\ell}|$. Now,

$$\sum_{j\geq 1}\frac{1}{j^{\frac{3}{2}}}E\left(\frac{1}{P_\ell(\theta,\xi)}\sum_{s=1}^{|Z^{(\ell)}|} e^{\theta z_{s,\ell}}|z_{s,\ell}|\right) \leq \sum_{j\geq 1} j^{-\frac{3}{2}}\int_{-\infty}^{\infty}|x|d\overline{H}_\ell(x)$$

$$\leq \sum_{j\geq 1} j^{-\frac{3}{2}}\int_{-\infty}^{\infty}|x|^2 d\overline{H}_\ell(x)$$

$$\leq K\sum_{j\geq 1} j^{-\frac{3}{2}}\ell \leq K\sum_{j\geq 1} j^{-\frac{3}{2}}j^{\alpha} < \infty.$$

(by the choice of α). Hence (4) tends to 0 as $n \to \infty$, completing the proof of the proposition.

Our next goal is to show that $(T_n - E(T_n|\mathcal{F}^{(\ell)}))$ converges to 0 with probability 1 for almost all environments. Conditionally on $\mathcal{F}^{(\ell)}$, T_n can be expressed as

$$T_n = \sqrt{2\pi B_n(\xi)}(P_n(\theta,\xi))^{-1}\sum_{s=1}^{|z^{(\ell)}}\sum_{r\in S_{s,\ell}} e^{\theta s^{z_{r,n}}f(s^{z_{r,n}})}$$

where $S_{s,\ell} = \{r\colon$ the position of the ancestor (in the ℓth generation) of a particle at $z_{r,n}$ (in the nth generation) is $z_{s,\ell}\}$ and ${}_s z_{r,n} = z_{r,n} + z_{s,\ell}$. The above expression can be rewritten as

$$T_n = \sqrt{2\pi B_n}(P_\ell(\theta,\xi))^{-1}\left(\sum_{s=1}^{|Z^{(\ell)}|} e^{\theta z_{s,\ell}}\right)\left[\left(P_{n-\ell}(\theta,T^\ell\xi)\sum_{r\in S_{s,\ell}} e^{\theta z_{r,n}} f({}_s z_{r,n})\right]\right.$$

The term inside the paranthesis can be expressed as

$$\int_{-\infty}^{\infty} f(x+z_{s,\ell})d_s\overline{Z}^{(n-\ell)}(x,T^\ell\xi)$$

where ${}_s\overline{Z}^{(n-1)}(\quad;T^\ell\xi)$ denotes the transformed point process of positions of the nth generation population starting with a single ancestor at $z_{s,\ell}$ in the ℓth generation. If $z_{s,\ell}=0$ we shall dispense with the s in the notation of the transformed point process.

Note also that ${}_s\overline{Z}^{n-\ell} \overset{d}{=} \overline{Z}^{(n-\ell)}$ where by $\overset{d}{=}$ we mean equality in the sense of distribution. Thus

$$\left(T_n - E(T_n|\mathcal{F}^{(\ell)})\right) = \sqrt{2\pi B_n(\xi)}\sum_{s=1}^{|Z^{(\ell)}|} a_s(\ell)Y_s(\ell) \text{ where}$$
$$a_s(\ell) = (P_\ell(\theta,\xi))^{-1}e^{\theta z_{s,\ell}}$$

and

$$Y_s(\ell) = \int_{-\infty}^{\infty} f(x+z_{s,\ell})d\left({}_s\overline{Z}^{(n-\ell)}(x,T^\ell\xi) - \overline{H}^{(n-\ell)}(x,T^\ell\xi)\right)$$

At this point we define a truncated version of $Y_{s,\ell}$ as follows:

$$Y_s^t(\ell) = \begin{cases} Y_s(\ell) & \text{if } |Y_s(\ell)| < (a_s(\ell))^{-1} \\ 0 & \text{if } |Y_s(\ell)| \geq (a_s(\ell))^{-1} \end{cases}$$

Thus,

$$\begin{aligned}\left(T_n - E(T_n|\mathcal{F}^{(\ell)})\right) = \sqrt{2\pi B_n(\xi)}\Bigg[\sum_{s=1}^{|Z^{(\ell)}|} a_s(\ell)E(Y_s^t(\ell)) &+ \sum_{s=1}^{|Z^{(\ell)}|} a_s(\ell)(Y_s - Y_s^t) \\ &+ \sum_{s=1}^{|Z^{(\ell)}|} a_s(\ell)(Y_s^t - E(Y_s^t))\Bigg] \qquad (5)\end{aligned}$$

We shall show that each of the terms converge to 0 with probability 1 (for almost all environments). The proofs will involve moment conditions on the tails of the distributions of $\sup_{n\geq 0} W^{(n)}(\theta, T^\ell\xi)$ for $\ell \geq 0$.

It is known in the branching processes theory that the distributions of the tails of $\sup_{n\geq 0} W^{(n)}(\theta, T^\ell\xi)$ are related to those of $W(\theta, T^\ell\xi)$ which in turn are related to those of $W^{(1)}(\theta, T^\ell\xi)$. Thus the moment conditions on $W^{(1)}(\theta, T^\ell\xi)$ transform to the finiteness of certain moments of $\sup_{n\geq 0} W^{(n)}(\theta, T^\ell\xi)$. The next two propositions make the above observations precise and the proofs are routine. The method is an adaptation of the one available for instance in Biggins (see [4]). We provide the proofs only to make the paper self-contained. We start with the following simple estimate

$$\begin{aligned}|Y_s(\ell)| &\leq K({}_sW^{(n-\ell)}(\theta, T^\ell\xi) + 1) \\ &\leq K(\sup_{n\geq 0} {}_sW^{(n)}(\theta, T^\ell\xi) + 1) \\ &\stackrel{d}{=} K(\sup_{n\geq 0} W^{(n)}(\theta, T^\ell\xi) + 1).\end{aligned}$$

Our next proposition provides the link between the moments of $W(\theta, T^\ell\xi)$ and $W^{(1)}(\theta, T^\ell\xi)$. Let $h\colon R^+ \to R^+$ be defined by

$$h(x) = \begin{cases} c_0 x & \text{if } x < e \\ c_1 + c_2 \log^{\frac{3}{2}} x & \text{if } x \geq e \end{cases}$$

where c_0, c_1, and c_2 are constants so chosen that h is concave. Since h is also non-decreasing it is sub-additive.

Proposition. *Under A1 - A10,* $E(W(\theta,T^{\ell}\xi)\log_{+}^{\frac{3}{2}} W(\theta,T^{\ell}\xi)) < \infty$.

Proof. It is enough to show that $E(W(\theta,T^{\ell}\xi)h(W(\theta,T^{\ell}\xi))) < \infty$ for almost all environments. We begin with

$$\begin{aligned}
W^{(n+1)}(\theta,T^{\ell}\xi) &= (P_{n+1}(\theta,T^{\ell}\xi))^{-1} \sum_{r=1}^{|Z^{(n+1)}|} e^{\theta z_{r,n+1}} \\
&= (P_n(\theta,T^{\ell}\xi))^{-1} \sum_{r=1}^{|Z^{(n)}|} e^{\theta z_{r,n}} {}_rW^{(1)}(\theta,T^{\ell+n}\xi) \\
&= (P_n(\theta,T^{\ell}\xi))^{-1} \sum_{r=1}^{|Z^{(n)}|} e^{\theta z_{r,n}} {}_rX_n(T^{\ell}\xi)
\end{aligned}$$

where ${}_rX_n(T^{\ell}\xi) \stackrel{d}{=} {}_rW^{(1)}(\theta,T^{\ell+n}\xi)$. Hence $W^{(n+1)}(\theta,T^{\ell}\xi)h(W^{(n+1)}(\theta,T^{\ell}\xi))$ is the same as

$$\left(\sum_{r=1}^{|Z^{(n)}|} \frac{e^{\theta z_{r,n}}}{(P_n(\theta,T^{\ell}\xi))} {}_rX_n(T^{\ell}\xi)\right) h\left(\sum_{s=1}^{|Z^{(n)}|} \frac{e^{\theta z_{r,n}}}{P_n(\theta,T^{\ell}\xi)} {}_rX_n(T^{\ell}\xi)\right).$$

The right hand side of the above equation is bounded above by

$$\begin{aligned}
&\left(\sum_{r=1}^{|Z^{(n)}|} \frac{e^{\theta z_{r,n}}}{P_n(\theta,T^{\ell}\xi)} {}_rX_n(T^{\ell}\xi)\right) h\left(\sum_{s\neq r} \frac{e^{\theta z_{r,n}}}{P_n(\theta,T^{\ell}\xi)} {}_rX_n(T^{\ell}\xi)\right) \\
&\quad + \sum_{r=1}^{|Z^{(n)}|} \frac{e^{\theta z_{r,n}}}{P_n(\theta,T^{\ell}\xi)} {}_rX_n(T^{\ell}\xi) h\left(\frac{e^{\theta z_{r,n}}}{P_n(\theta,T^{\ell}\xi)} {}_rX_n(T^{\ell}\xi)\right)
\end{aligned}$$

Now taking conditional expectations with respect to $\mathcal{F}^{(n)}$, we have

$$\begin{aligned}
E(W^{(n+1)}(\theta,T^{\ell}\xi)h(W^{(n+1)}(\theta,T^{\ell}\xi))|\mathcal{F}^{(n)}) &\leq W^{(n)}(\theta,T^{\ell}\xi)h(W^{(n)}(\theta,T^{\ell}\xi)) \\
&\quad + X_n(T^{\ell}\xi) \int_{-\infty}^{\infty} h\left(\frac{e^{\theta x} X_n(T^{\ell}\xi)}{P_n(\theta,T^{\ell}\xi)}\right) d\overline{Z}^{(n)}(x,T^{\ell}\xi)
\end{aligned}$$

Taking expectations and iterating the above inequality, one has, using Fatou's lemma,

$$E\left(W(\theta,T^\ell\xi)h(W(\theta,T^\ell\xi))\right) \le E\left(\sum_{n\ge 0} X_n(T^\ell\xi)\int_{-\infty}^{\infty} h\left(\frac{e^{\theta x}X_n(T^\ell\xi)}{P_n(\theta,T^\ell\xi)}d\overline{H}^{(n)}(x,T^\ell\xi)\right)\right).$$

To complete the proof of the theorem, all we need to do is establish the finiteness of the right hand side. Consider

$$E_{\xi_0}\left(E(\sum_{n\ge 0} X_0(T^{\ell+n}\xi)\int_{-\infty}^{\infty} h\left(\frac{e^{\theta x}X_0(T^{\ell+n}\xi)}{P_n(\theta,T^\ell\xi)}d\overline{H}^{(n)}(x,T^\ell\xi 0\right)\right)$$

which is the same as (by Tonelli's Theorem and monotone convergence theorem)

$$E_{\xi_0}\left(E\left(\sum_{n\ge 0} X_0(\xi)\int_{-\infty}^{\infty} h\left(\frac{e^{\theta x}X_0(\xi)}{P_n(\theta,\xi)}\right)d\overline{H}^{(n)}(x,\xi)\right)\right).$$

Next consider

$$\sum_{n\ge 0}\int_{-\infty}^{\infty} h\left(\frac{e^{\theta x}X_0(\xi)}{P_n(\theta,\xi)}\right)d\overline{H}^{(n)}(x,\xi)$$

which can be rewritten as

$$\sum_{n\ge 0}\int_{I_n} h\left(\frac{e^{\theta x}X_0(\xi)}{P_n(\theta,xi)}\right)d\overline{H}^{(n)}(x,\xi)+\sum_{n\ge 0}\int_{I_n^c} h\left(\frac{e^{\theta x}X_0(\xi)}{P_n(\theta,\xi)}\right)d\overline{H}^{(n)}(x,\xi).$$

We shall first give a choice for I_n. Since under our assumptions $\inf_{k\ge 0} m_k(\theta) > 1$ for all $k > 0$ for almost all environments, for every $n \ge 0$ there exists a w such that

$$\inf\{e^{-\theta x}P_n(\theta):\ -nw \le x \le wn\} \ge \beta^n > 1.$$

choose $I_n^c = [-wn, wn]$. For such a choice of I_n, one has

$$\sum_{n\ge 0}\int_{I_n^c} h\left(\frac{e^{\theta x}X_0(\xi)}{P_n(\theta,\xi)}\right)d\overline{H}_n(x,\xi) \le \sum_{n\ge 0} h\left(X_0(\xi)\beta^{-n}\right)$$

and on I_n, using the estimate

$$h(xy) \leq K(1 + \log_+^{\frac{3}{2}} x + \log_+^{\frac{3}{2}} y)$$

we have

$$\sum_{n\geq 0} \int_{I_n} h\left(\frac{e^{\theta x} X_0(\xi)}{P_n(\theta,\xi)}\right) d\overline{H}_n(x,\xi)$$

is bounded above by

$$\sum_{n\geq 0} \int_{I_n} (1 + \log_+^{\frac{3}{2}} X_0(\xi) + \log_+^{\frac{3}{2}} \left(\frac{e^{\theta x}}{P_n(\theta,\xi)}\right) d\overline{H}_n(x;\xi)$$

which in turn is bounded by

$$\sum_{n\geq 0} \int_{I_n} (1 + \log_+^{\frac{3}{2}} X_0(\xi) + \log_+^{\frac{3}{2}}(\theta x)) d\overline{H}^{(n)}(x,\xi).$$

We shall first show that $E_{\xi_0}\left(E\left(\sum_{n\geq 0} h(X_0(\xi)\beta^{-n})\right)\right) < \infty$. Choosing $N = \left[(\log\beta)^{-1}\log\left(\frac{X_0(\xi)}{x_0}\right)\right]$,

$$\begin{aligned}
\text{we see that, } \sum_{n\geq 1} h(X_0(\xi)\beta^{-n}) &\leq h(X_0(\xi))N + C_0 \sum_{n\geq N+1} X_0(\xi)\beta^{-n} \\
&\leq \frac{h(X_0(\xi))}{\log\beta}\log\left(\frac{X_0(\xi)}{X_0}\right) + C_0 x_0 \\
&\leq K(1 + \log_+^{\frac{5}{2}} X_0(\xi))
\end{aligned}$$

and hence the result follows from A7.

As for the other terms, we shall only estabish that $\sum_{n\geq 0} \int_{nw}^{\infty} x^{\frac{3}{2}} d\overline{H}^{(n)}(x,\xi) < \infty$

Note that

$$\sum_{n\geq 0}\int_{nw}^{\infty} x^{\frac{3}{2}}\, d\overline{H}^{(n)}(x,\xi) < \infty = \sum_{n\geq 0}\sum_{r\geq n}\int_{rw}^{(r+1)w} x^{\frac{3}{2}} + x d\overline{H}^{(n)}(x)$$

$$\leq K\sum_{n\geq 0}\sum_{r\geq n}(r+1)^{\frac{3}{2}}\overline{H}^{(n)}((r+1)w)$$

$$\leq K\sum_{r\geq 0}(r+1)^{\frac{3}{2}}e^{-(r+1)w\eta}\sum_{r=0}^{r}\prod_{k=0}^{n}\frac{m_k(\theta+\phi)}{m_k(\theta)}$$

$$\leq K\sum_{r\geq 0}(r+1)^{\frac{3}{2}}c^r$$

where $0 < c < 1$ and K is independent of the environment; hence it follows that $E_{\xi_0}(E(W(\theta,\xi)\log W^{\frac{3}{2}}(\theta,\xi))) < \infty$ from which it also follows that $E(W(\theta,\xi)\log W^{\frac{3}{2}}(\theta,\xi))$ is finite with probability 1 for almost all environments.

Our next proposition relates the tail of distributions of $W(\theta, T^{\ell}\xi)$ with that of $\sup_{n\geq 0} W^{(n)}(\theta, T^{\ell}\xi)$ and the proof follows from the proof given in Biggins (see [4]) and hence will be omitted.

Proposition 4.3. *Under A, for almost all environments and for every $0 < a < 1$ and $\ell \geq 0$, there exists $B > 0$ such that*

$$P(W(\theta,T^{\ell}\xi) \geq at) \geq BP(\sup_{n\geq 0} W^{(n)}(\theta,T^{\ell}\xi) \geq t) \geq BP(\sup_{n\geq 0} W(\theta,T^{\ell}\xi) \geq t).$$

Thus to complete the proof of the theorem we need to show the convergence of the three terms in the RHS of (5). The proof of convergence of the second and third terms to zero follow from the arguments in Biggins [4]. The proof of the convergence to zero of the first term is contained in Proposition 4.4.

Proposition 4.4. $\sqrt{2\pi B_n(\xi)}\sum_{s=1}^{|Z^{\ell}|} a_s(\ell)E(Y_s^t(\ell))$ *converges to 0 with probability 1 for almost all environments.*

Proof. First observe that

$$E(Y_s^t(\ell)) = E(Y_s(\ell)\colon |Y_s(\ell)| < (a_s(\ell))^{-1})$$
$$= E(Y_s(\ell)\colon |Y_s(\ell)| > (a_s(\ell))^{-1}).$$

Also by ergodic theorem, there exists $N_0(\xi)$ such that for all $n \geq N_0, B_n \leq Kn$ for some finite positive constant K (possibly depending on $\overline{\xi}$). Thus,

$$\sqrt{2\pi B_n(\xi)} \sum_{s=1}^{|Z^{(\ell)}|} a_s(\ell) E(Y_s^t(\ell)) \leq K\sqrt{n} \sum_s a_s(\ell) \int_{(a_s(\ell))^{-1}}^{\infty} t dG_\ell(t)$$
$$\leq K(j+1)^{\frac{3}{2}} \sum_s a_s(\ell) \int_{(a_s(\ell))^{-1}}^{\infty} t dG_\ell(t).$$

Let $R = \{s\colon a_s(\ell) > e^{-m\ell}\}$ where $0 < m < K(\theta)$. Then,

$$\sum_s a_s(\ell) \int_{(a_s(\ell))^{-1}}^{\infty} t dG_\ell(t) = \left(\sum_{s\in R} + \sum_{s\in R^c}\right) \left(a_s(\ell) \int_{(a_s(\ell))^{-1}} t dG_\ell(t)\right)$$
$$\leq \left(\sum_{s\in R} a_s(\ell)\right) \int_0^{\infty} t dG_\ell(t) + W^{(\ell)}(\theta,\xi) \int_{e^{m\ell}}^{\infty} t dG_\ell(\theta)$$

Now,

$$E\left(\sum_j (j+1)^{\frac{3}{2}} \sum_{s\in R} a_s(\ell)\right) \leq \sum_j (j+1)^{\frac{3}{2}} E\left(\sum_{s\in R} a_s(\ell)\right)$$
$$\leq \sum_j (j+1)^{3/2} c^{j^\alpha} < \infty$$

Also,

$$(j+1)^{\frac{3}{2}} W^{(\ell)}(\theta,\xi) \int_{e^{m\ell}}^{\infty} t dG_\ell(t) \leq K W^{\ell}(\theta,\xi) \int_{e^{m\ell}}^{\infty} t (\log t)^{\frac{3}{2}} dG_{\ell)(t)}$$

$\to 0$ as $j \to \infty$, thus completing the proof of the proposition.

5. References.

1. Asmussen and Kaplan (1976): *Branching Random Walks I and II*, Journal of Stochastic Processes and Applications **4**, 1-32.
2. Athreya and Karlin (1972): *On Branching Processes in random environments: I, Extinction Probability*, Annals of Mathematical Statistics **42**, 1499-1520.
3. Athreya and Karlin (1972): *On branching processes in random environments: II, Limit theorems*, Annals of Mathematical Statistics **F2**, 1843-1858.
4. Biggins, J. D. (1979): *Growth rates in the branching random walk*, Z. Wahrsch. Verw. Gebiete **48**, 17-34.
5. Biggins, J. D. (1989): *Uniform convergence of martingales in the one-dimensional branching random walk*, In selected proceedings of the symposium on Applied Probability, Sheffield (I. V. Basawa and R. L. Taylor, eds.), vol. 18, IMS, Hayward, California, pp. 159-173.
6. Bramson, M., Ney, P., and Tao, J. (1992): *Population composition of a multitype branching random walk*, Annals of Applied Probability **2, No. 3**, 575-596.
7. Kallenberg (1983): *Random Measures*, Academic Press, New York.
8. Petrov (1975) *Sum of independent random variables*, Springer-Verlag, Berlin.
9. Tanny, D. (1977): *Limit theorems for branching processes in random environment*, Annals of Probability **5**, 100-116.
10. Tanny D. (1988): *A necessary and sufficient conditioon for a branching process in a random environment to grow like the product of its means*, Stochastic Processes and their Applications **28**, 123-139.
11. Vidyashankar, A. N. (1994): *Kesten-Stigum type theorem for BRWRE*, submitted.

GENERAL SUMMARY

In this thesis we investigated the large deviation problems for branching models by focussing on three distinct questions. First we investigated the rates of decay for the deviation between empirical mean and the true mean for various branching processes. We showed that (under further moment and regularity conditions) in the discrete time case the rate is geometric while in the continuous time case the rate is exponential.

Second, we studied the large deviation behavior exhibited by the tails of the martingale limit viz. W. More precisely, we studied the behavior of $P(W \leq x)$ near zero by obtaining bounds on the density w of W near the origin and then transformed these bounds to extract the rate of decrease of the left tail. A related problem that we investigated concerned the right tail of W. We showed that if the branching process has finite support then one can use the multitype analogues of the Harris function to obtain precise information about the behavior of $P(W \geq x)$ for large x.

Finally, we discussed the branching random walk model in stationary ergodic environments. We showed that (conditioned on the environment) the number of particles living to the left of nx (for a certain x) normalized by its expectation converges to a non-negative random variable with probability one which is non-degenerate.

There are several related questions that are as yet unresolved. We shall mention a few of them:

(1) In the single type Galton-Watson case, the non-parametric maximum likelihood estimator of the mean m is

$$Y_n = \frac{\sum_{j=0}^{n} Z_j}{\sum_{j=0}^{(n-1)} Z_j}.$$

It can be shown that $Y_n \to m$ with probability one. Since Y_n can be expressed as $\sum_{j=1}^{n-1} a_j T_j$ where $T_j = \frac{Z_j}{Z_{j-1}}$ and $a_j = \frac{Z_j}{\sum_{k=0}^{n-1} Z_k}$ it is tempting to conjecture that $P(|Y_n - m| > \epsilon) \sim p_1^n$. Is such a conjecture true? Similar questions for multi-type cases will also be interesting.

(2) In the multi-type case, we showed that if the matrix $A = 0$, then under the assumption that every particle produces at least of its kind that

$$-log f_i^{(n)}(s_1, s_2) \sim 2^n R_i(s_1, s_2) \quad \text{for } i = 1, 2.$$

Is this result true without these conditions?

(3) In the braching random walk in random environments we obtained a necessary and sufficient condition for the non-degeneracy of a certain martingale limit $W(\theta)$ for θ belonging to a certain set. What would happen if θ were not in such a set? At the present time we do not know how to examine questions of this type and the problem seems to be challenging.

LITERATURE CITED

1. Athreya, K.B., (1994): *Large deviation rates for branching processes-1, the single type case*, to appear in Annals of Applied Probability.
2. Biggins, J.D. and Bingham, N.H., (1993): *Large deviations in the supercritical branching processes*, Advances in Applied Probability, **25**, #4, 757– 772.
3. Harris, (1948): *Branching Processes*, Ann. Math. Statist., **41**, 474–494.
4. Karp, R. and Zhang, Y., (1993): *Tail probabilities for finite supercritical branching processes*, Technical Report, Department of Computer Science and Engineering, Southern Methodist University, Dallas, Texas.
5. Miller, H.I. and O'Sullivan, J.A., (1992): *Entropies and combinatorics of random branching processes and context free languages*, IEEE transactions on information theory 38, #4, July, 1292–1311.

ACKNOWLEDGEMENTS

Many people helped me during my study here at Iowa State University. Though I cannot name all of them here, I would like to mention a few specific people who made my stay at ISU memorable. A long distance role was played by my mother who encouraged me to excel in everything I did. I would like to express my deep sense of gratitude to her.

Professionally, I had an excellent relationship and friendship with my thesis advisor Professor Krishna B. Athreya. We worked long on several problems and he gave me lots of his time and advice. I thank him for all of his help.

Doug Betsinger and Dale Buske helped me with my chores during my visits to several meetings and conferences. I thank them for all of their help. Most of this thesis was typed by Sally Emmerson, who deserves a very special thanks. But for her help, this would not have been easy. Jan, Ruth, and Ellen helped me on several occasions during my study here. I would like to express my thanks to them.

Finally, I would like to thank Professor Wayne Fuller of the Statistics Department and the Mathematics Department for supporting my studies from 1988 - 1994.